Titre: Title:	3D reconstruction of coronary arteries from angiographic sequences for interventional assistance
Auteur: Author:	Pascal Fallavollita
Date:	2008
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Fallavollita, P. (2008). 3D reconstruction of coronary arteries from angiographic sequences for interventional assistance [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/8149/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/8149/
Directeurs de recherche: Advisors:	
Programme: Program:	Unspecified

UNIVERSITÉ DE MONTRÉAL

3D RECONSTRUCTION OF CORONARY ARTERIES FROM ANGIOGRAPHIC SEQUENCES FOR INTERVENTIONAL ASSISTANCE

PASCAL FALLAVOLLITA
INSTITUT DE GÉNIE BIOMÉDICAL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(GÉNIE BIOMÉDICAL)
JUILLET 2008

Your file Votre référence
ISBN: 978-0-494-46100-6
Our file Notre référence
ISBN: 978-0-494-46100-6

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

3D RECONSTRUCTION OF CORONARY ARTERIES FROM ANGIOGRAPHIC SEQUENCES FOR INTERVENTIONAL ASSISTANCE

présentée par: FALLAVOLLITA Pascal

en vue de l'obtention du diplôme de: Philosophiæ Doctor

a été dûment acceptée par le jury d'examen constitué de:

M. SAVARD Pierre, Ph.D., président

Mme. CHERIET Farida, Ph.D., membre et directrice de recherche

M. DELORME Sébastien, Ph.D., membre

Mme. NOUMEIR Rita, Ph.D., membre

If we knew what it was we were doing, it would not be called research, would it?

~Albert Einstein

ACKNOWLEDGEMENTS

To my supervisor, Professor Farida Cheriet, I owe my everlasting gratitude for having guided me in these last few years. My research work is directly related to her constant support and willingness to see me succeed. But above all, she provided me the freedom to pursue my research independently and her availability and capacity to answer all questions are quite fascinating to this day.

I am thankful to Dr. Chenyang Xu, my internship supervisor at Siemens Corporate Research, Princeton, NJ, for passing on his expertise and passion in order for me to complete my project with them. The positive experience at SCR has been an unforgettable one.

I thank Professor Guillaume Alexandre Bilodeau, who gave me the opportunity to work as a teaching assistant in his classes these last few years. A huge thanks to Professor Pierre Savard, who gave me my first opportunity at becoming a young investigator in the biomedical field.

Thank you to Rafik Bourezak (*in Algeria somewhere!*), Thomas Hurtut (*how are you?*), Rola Harmouche (*bru-schetta!*), Lyes Djerroud (*I'm Canadian!*) and every other LIV4D colleagues – thank you for always being open to discussions, whether technical or not, and providing so many worthy lessons that cannot be found in textbooks. More particularly to Ionut Alexandrescu, who has been there every step of the way in providing valuable programming advice.

Thank you also to the Natural Sciences and Engineering Research Council of Canada, whose funding supports the project that this thesis contributes toward.

Lastly, to my parents for their sacrifice and love, it could not have happened without you. And to Ann-Marie, who sacrificed more than I did in putting up with me these last few years and for her encouragement, this work could not have happened without you.

RÉSUMÉ

Les maladies cardiovasculaires sont la première cause de décès en Amérique du nord. Ces maladies impliquent généralement des dépôts lipidiques et de calcium sur les parois internes des artères coronariennes entraînant ainsi une des plus communes maladies connue sous le nom d'athérosclérose. Une plaque d'athérome se fixe progressivement sur les parois internes des artères et peut entraîner un blocage partiel ou total du flux sanguin à travers l'artère. L'angiographie par rayons-X est la technique de première ligne pour le diagnostic et le traitement des anomalies morphologiques du réseau coronarien.

Les différentes institutions hospitalières sont souvent équipées de systèmes monoplans ce qui a conduit à une acquisition séquentielle de deux séquences monoplans synchronisées sur l'électrocardiogramme pour permettre l'application des techniques de reconstruction 3D biplan. Ainsi, la précision des reconstructions 3D obtenues reste limitée à cause des inévitables artefacts de mouvement dus aux acquisitions non simultanées des séquences d'images. Cependant, les procédures d'angioplastie se font en utilisant seulement une vue en maintenant un angle d'incidence fixe tout au long de l'intervention.

L'objectif général de la programmation de recherche proposée est de développer et valider des algorithmes innovateurs qui permettront de concevoir un outil clinique pour assister les cardiologues pendant les procédures d'angioplastie. La technologie principale utilisée dans les salles fluoroscopiques est l'imagerie à rayons-X, donc il est impératif d'exploiter toute l'information disponible contenue dans les images angiographiques afin de faciliter l'intervention du cardiologue et de réduire le temps de diagnostique et de décision pendant les chirurgies. Afin de répondre à cet objectif, deux hypothèses de recherche ont été formulées. La première consiste à reconstruire l'artère coronaire en 3D. Une reconstruction fiable est directement liée aux points de correspondance choisis et à la robustesse de leur appariement entre les deux vues. Et la

deuxième consiste à formuler des concepts pour développer un algorithme monoplanaire pour la reconstruction 3D des artères.

Lors d'une première étape, l'image angiographique représentant la diastole est extraite en calculant l'écart-type de la différence entre deux images consécutives. Deuxièmement, l'image diastolique est filtrée par quatre filtres différents: homomorphique, anisotropique, choc et morphologique. Ces filtres sont utilisés pour la restauration et le rehaussement des images angiographiques. En effet, les images angiographiques sont fortement bruitées et un flou local est souvent présent à cause des artefacts de mouvement et des variations d'intensité dues à l'avancement de l'agent de contraste dans le sang. Étant intéressés par la reconstruction de la ligne centrale des vaisseaux ou des cathéters utilisés lors des interventions, nous avons utilisé une méthode du *Fast Marching Method (FMM).* Cette méthode offre à l'utilisateur la possibilité de sélectionner deux points sur une artère spécifique et d'extraire automatiquement la ligne centrale qui relie ces deux.

Le suivi automatique de la ligne centrale à travers la séquence d'images est effectué en utilisant une procédure basée sur le flot optique multi-échelle et des contours actifs *Gradient Vector Flow (GVF).* Le flux optique multi-échelle permet de calculer les mouvements d'artères sachant qu'elles peuvent effectuer de grands déplacements entre deux images consécutives. Ces estimations de déplacements en pixels, sont additionnées aux positions de la ligne centrale dans la première image pour obtenir les coordonnées approximatives de la ligne centrale dans la deuxième image. Ces coordonnées sont utilisées pour paramétrer le contour actif. Une fois les lignes centrales des structures d'intérêts segmentées et suivies sur une séquence monoplan d'images angiographiques, une procédure d'auto-calibrage est développée afin de reconstruire ces lignes centrales en 3D. Nous ajoutons une contrainte sur la courbure 2D/3D des artères pour filtrer les correspondances trouvées par la méthode géométrique épipolaire. Le choix de la procédure d'auto-calibrage monoplanaire a été fortement motivée par le fait suivant: il est techniquement faisable de déterminer les paramètres intrinsèques et extrinsèques d'un système à rayons-X à partir du contenu naturel des images si un modèle 3D *a priori*

des artères à un instant donné est disponible. De plus, il a été clairement démontré que le contenu temporel des séquences peut être exploité afin de compléter l'information spatiale.

Pour répondre à la première hypothèse, en ajoutant une contrainte d'optimisation de la procédure de correspondance, qui spécifie la relation entre la courbure 2D des repères dans les images et la courbure de Frenet-Serret en 3D, les résultats de reconstruction ont été améliorés. En utilisant cinq paires d'images consécutives, pour une configuration antéro-postérieur l'erreur 3D RMS a diminué de 2.8mm à 1.1mm, et pour une configuration oblique antérieure droite/oblique antérieure gauche l'erreur a diminué de 3.1mm à 1.9mm.

Pour répondre à la deuxième hypothèse, nous proposons la stratégie suivante pour la reconstruction 3D monoplanaire des artères (ou cathéter). En supposant que nous disposions du modèle 3D de la structure d'intérêt à un instant donné, il est possible d'estimer la géométrie 3D de la structure en utilisant un minimum de trois images à rayons-X consécutives. De plus, la méthodologie devient plus précise en exploitant les sections de structures comportant des mouvements strictement rigides, tel les bifurcations ou électrodes du cathéter. Une étude clinique a été effectuée pour l'estimation de la profondeur d'une électrode sur trois images consécutives. Pour une configuration antéro-postérieur l'erreur 3D moyenne et maximale de la profondeur est de 2.68mm et 7.05mm respectivement, et pour une configuration oblique antérieure droite/oblique antérieure gauche l'erreur moyenne est de 2.34mm et l'erreur maximale est de 5.84mm. De futures études doivent mettre l'emphase sur des contraintes de non-rigidité pour tenir compte du mouvement local du cœur, des structures anatomiques et des instruments.

La recherche proposée ici établira les bases scientifiques d'outils de visualisation qui visent à faciliter l'intervention et d'assister les cardiologues dans ces procédures. Deux hypothèses ont été vérifiées. La première hypothèse a permis de montrer qu'une contrainte géométrique a permis de raffiner toutes les correspondances possibles pour la reconstruction, et par conséquent, a amélioré la précision de cette reconstruction. On

vise des segments d'artères ou des cathéters qui subissent seulement des mouvements rigides afin d'exploiter la non-contractilité de ces structures et d'assurer un appariement plus robuste en considérant la courbure en 2D et en 3D. La deuxième hypothèse a permi de démontrer qu'il y a possibilité de reconstruire des structures en 3D en utilisant seulement une séquence angiographique et un modèle 3D *a priori* à un instant donné.

ABSTRACT

Cardiovascular disease is still the leading cause of death in North America today. One of its forms, atherosclerosis, is manifested from the build-up of plaque inside the coronary artery walls which may lead to stenosis, a partial or complete blockage inside the artery. In the most urgent cases, an angioplasty procedure is performed in order to crush the plaque, using a balloon catheter, or reestablish normal blood flow by fixating a stent inside the artery.

Angioplasty procedures are performed in sterile rooms called fluoroscopy labs. X-ray fluoroscopic imaging remains the method of choice of cardiologists for proper coronary heart disease diagnosis and for interventional procedures. Acquiring a pair of monoplane X-ray fluoroscopic images, over a cardiac phase, facilitates three dimensional reconstruction of the coronary arteries and enables the cardiologist to look at the volume and make a proper diagnostic. However, in most cases, the C-arm fluoroscope is fixed at a specific inclination angle during the procedure for intervention guidance purposes.

The main objective of this work is to implement novel techniques necessary to extract and use the information contained in single plane angiographic sequences in order to develop a potential tool to assist the cardiologist concurrently during the procedure. Two specific hypotheses will need to be verified to justify our main research objectives. The first hypothesis states that the integration of a geometrical curvature constraint improves the precision of the point correspondence and improves reconstruction accuracy whereas the second hypothesis states that it is possible to recuperate the 3D geometry across a single plain angiographic sequence, for clinical assistance, by using an *a priori* 3D model of the structure of interest.

We implemented a novel automatic 2D segmentation algorithm to enhance healthy or diseased (stenosis) coronary arteries in a first image representing the diastolic cardiac phase. A 4-step filter, that included a homomorphic, anisotropic, shock filter and morphological component systematically suppressed the background and enhanced only the arteries. Then, we

provide the cardiologist with the ability to target a specific coronary artery and investigate its motion by extracting its 2D centerline and automatically temporal tracking it across all the angiographic images in a cardiac cycle. A two click Fast Marching Method for the centerline extraction phase, whereas a gradient vector flow (GVF) active contour model coupled with a pyramidal Optical Flow approach is used for the temporal tracking procedure. These tools lead towards the reconstruction process of a targeted coronary artery. Valid 3D reconstruction relies on the successful correspondence of appropriate landmarks obtained from the arteries through a set of angiographic images. By introducing a novel curvature constraint that relates 2D-3D coronary curvature, we aim at reconstructing structures that undergo mainly rigid motion, since if this is the case, the structure would not undergo contractility and hence would have similar curvature between its 3D geometry and its projected shape in the 2D images. The proposed algorithms were validated by using 5 consecutive biplane image pairs, we obtain the following research results: for a posterior/anterior view, the 3D RMS improves from 2.8mm to 1.1mm, whereas for a left-right anterior oblique setup the 3D RMS improves from 3.1mm to 1.9mm using our geometric curvature constraint.

Recalling that the cardiologist usually makes use of a single plane angiographic view by looking at the X-ray fluoroscopy monitor during interventions, it would be advantageous to reconstruct the arteries in 3D using one view. Thus, we suppose that an *a priori* 3D model of the arteries is obtained and available at a first time instant. The model can be obtained from a biplane reconstruction before the intervention, or from preoperative data, such as an MRI or MDCT dataset, which are the registered thereafter in the coordinate frame of the fluoroscope during a procedure. The monoplane equations developed optimize for the unknown 3D displacements in subsequent time instants of the angiographic dataset. A clinical experiment was performed using arrhythmia data. The aim is to estimate the depth of the tip electrode of the ablation catheter in 20 distinct datasets. For a posterior/anterior view, the 3D mean and maximum depth errors were 2.68mm and 7.05mm, whereas for a left-right anterior oblique setup the 3D mean and maximal errors were 2.34mm and 5.84mm respectively.

In conclusion, our research work helped verify the two hypotheses outlined above. In order to develop a potential tool to assist the cardiologist during angioplasty procedures primarily (i.e. catheter ablation procedures may be targeted as well), we developed solutions that consider and extract the spatio-temporal information from X-ray images acquired by the fluoroscope. We implemented a two-click process that extracts the coronary artery centerline in the diastolic image and that is tracked temporally using all images representing an entire cardiac cycle. The 3D reconstruction obtained using a biplane setup was optimized by considering a novel geometric curvature constraint when choosing point correspondences. Single plane reconstruction is made possible if an *a priori* 3D model of the targeted structure is available at a 1st time instant in the angiographic images. Subsequent estimates of the 3D structure are obtained by considering the spatial coordinates and displacements of the arteries as well. Future work should focus on including non-rigid constraint equations to consider the inherent movement of the heart. Other geometrical constraints such as length preservation terms and 3D perpendicularity constraints between normals and tangents could be exploited to improve results using a single view.

CONDENSÉ

Introduction

Les maladies cardiovasculaires sont la première cause de décès en Amérique du nord. Environ deux fois plus de personnes décèdent d'une telle maladie que de toutes les autres formes de cancer. Ces maladies impliquent généralement des dépôts lipidiques et de calcium sur les parois internes des artères coronariennes entraînant ainsi une des plus communes maladies connue sous le nom d'athérosclérose. Une plaque d'athérome se fixe progressivement sur les parois internes des artères et peut entraîner un blocage partiel ou total du flux sanguin à travers l'artère.

Durant les 30 dernières années, l'angiographie par rayons-X a été la technique de première ligne pour le diagnostic et le traitement des anomalies morphologiques du réseau coronarien. Les images 2D obtenues sont analysées manuellement ou automatiquement afin de mesurer la sévérité de cette maladie, qui est aussi appelée sténose, à partir du rétrécissement des vaisseaux entraîné par la formation de la plaque. Cependant, les mesures obtenues sont intimement liées à l'angle d'incidence de la vue obtenue sachant que la plaque ne se forme pas de façon symétrique sur les parois internes. En effet, une sous-estimation de la sévérité de la maladie est souvent obtenue car un vaisseau peut paraître sain selon une vue bien que souffrant d'une sténose dont la sévérité peut atteindre jusqu'à 50%.

Durant ces dernières années, la disponibilité des systèmes d'angiographie biplan a permis la reconstruction 3D de l'arbre coronarien pendant quelques cycles cardiaques. Ceci permet une évaluation plus objective de la sévérité de la sténose. Par contre, la procédure de calibrage requise pour la reconstruction 3D se base sur les points de bifurcations car ils sont plus faciles à mettre en correspondance sur les deux images. En revanche, le nombre de points de bifurcations peut être limité dans les régions d'intérêt ce qui implique la nécessité d'extraire d'autres points de correspondance des artères coronaires. D'autre part, le calibrage du système biplan est essentiel pour effectuer une reconstruction précise. La procédure de calibrage permet d'estimer les paramètres géométriques des sources à rayons-X tels que les angles d'inclinaison, la distance focale

et les points principaux. Cependant, les différentes institutions hospitalières sont souvent équipées de systèmes monoplans ce qui conduit à une acquisition séquentielle de deux séquences monoplans synchronisées sur l'électrocardiogramme pour permettre l'application des techniques de reconstruction 3D biplan. Ainsi, la précision des reconstructions 3D obtenues reste limitée à cause des inévitables artefacts de mouvement dus aux acquisitions non simultanées des séquences d'images.

L'objectif général du projet de thèse est de développer et valider des algorithmes innovateurs qui permettront de développer un outil clinique pour assister les cardiologues pendant les procédures d'angioplastie. L'outil clinique envisagé permettra aux cardiologues de visualiser en 2D et en 3D les artères coronaires pendant la chirurgie. Sachant que la modalité principale utilisée dans les salles fluoroscopiques est l'imagerie à rayons-X, il est impératif d'exploiter toute l'information disponible dans les images angiographiques afin de faciliter l'intervention du cardiologue et de réduire le temps de diagnostique et de décision pendant les chirurgies.

Les deux contributions principales de ce projet de thèse sont liées directement à l'obtention d'une reconstruction 3D de l'artère coronaire. Premièrement, un calibrage fiable et précis du système d'acquisition d'images est directement associé aux points de correspondance choisis et à la robustesse de leur appariement entre les deux vues. On propose donc un algorithme optimal et innovateur pour raffiner et choisir seulement les points de correspondances fiables entre deux images angiographiques. Ainsi, un calibrage précis du système permettra une reconstruction 3D plus précise des structures d'intérêt à partir d'une paire de vues non simultanées. Deuxièmement, la contribution majeure de ce projet et de fournir une reconstruction 3D des artères coronaires aux cardiologues en utilisant seulement une vue angiographique monoplanaire acquise durant l'intervention et un modèle 3D des structures obtenu avant l'intervention soit par reconstruction 3D à partir d'une paire de vues non simultanées soit à partir d'une autre modalité d'imagerie. À notre connaissance, aucune technique de ce genre n'a été présentement proposé dans la communauté scientifique.

Pour réaliser ces deux contributions principales 4 objectifs spécifiques à considérer sont:

1. Développement d'un filtre permettant de rehausser les contours des artères angiographiques tout en supprimant l'arrière plan de l'image.
2. Conception d'un algorithme semi-automatique permettant à l'utilisateur de sélectionner deux points de repères sur une artère spécifique et d'extraire sa ligne centrale automatiquement. Ceci permettra le développement d'un algorithme automatique de contour actif permettant de suivre temporellement une artère coronaire à travers un cycle cardiaque.
3. Intégration d'une contrainte 2D/3D permettant le raffinement du processus de sélection de points correspondants sur deux vues non simultanées, pour la reconstruction 3D optimale des artères.
4. Effectuer une reconstruction monoplanaire de structures rigides présentes dans des images angiographiques. Les structures rigides sont ciblées car on vise des structures qui ne se contractent pas, et donc, qui auront une courbure semblable dans les images et leurs projections en 3D.

État de l'art

Techniques de rehaussement d'images

Appliqués à l'analyse d'images, les filtres de Lorenz et al. [6] et Frangi et al. [7] utilisent la matrice Hessienne permettant de décrire les variations locales d'intensités autour d'un point spécifique. La matrice Hessienne est un tenseur composé des sorties d'opérateurs dérivatifs du deuxième ordre. Une des difficultés de l'utilisation des dérivées secondes est leur forte sensibilité au bruit présent dans les images angiographiques. C'est pourquoi une convolution Gaussienne est combinée avec la deuxième dérivée de la matrice Hessienne pour effectuer un lissage de l'image et ainsi diminuer le niveau du bruit. L'orientation locale est donnée par le minimum de la dérivée seconde. Ce problème se ramène à la diagonalisation de la matrice Hessienne. Chaque valeur propre indique alors la variation locale d'intensité selon le vecteur propre correspondant. Une des limites de cette approche est que l'utilisateur doit sélectionner judicieusement l'échelle σ, appropriée pour le calcul des dérivées secondes. Plus l'échelle est grande, plus le

rehaussement des artères est prononcé au détriment du rehaussement des petits vaisseaux. Lorenz utilise seulement la plus grande valeur propre pour le rehaussement d'images tandis que Frangi va considérer les deux premières valeurs propres.

Contours actifs

La représentation implicite des contours actifs (les « level sets »), présente moins de contraintes quant à la géométrie des objets qui peuvent être irréguliers et dont la topologie peut changer en cours de convergence. Caselles et al. [24] ont introduit un modèle géométrique des contours actifs formulé par des équations aux dérivées partielles (EDPs) d'évolution de courbes. Chaque point de la courbe se déplace dans la direction de son vecteur normal à une vitesse proportionnelle à la courbure. Cette approche est une alternative à la minimisation d'énergie qui peut être vue comme la recherche de la solution d'une équation de Hamilton-Jacobi, et être ainsi résolue de façon efficace par l'utilisation de la méthode des courbes de niveaux d'Osher [26] et de Sethian [27]. L'utilisation de cette représentation implicite explique tout le succès de cette approche.

Les contours actifs explicites utilisent classiquement l'information de courbure pour compléter la mesure ambiguë du gradient d'intensité en chaque point de l'image, en pénalisant le développement de courbures élevées dans le contour actif. Une telle courbure peut par exemple se développer autour du bruit local, créant des gradients forts mais ne correspondant pas aux frontières d'un objet d'intérêt. Malheureusement, cette approche est légèrement limitée. Le contour actif doit en effet être paramétrisé par un ensemble de particules devant être supprimées, rajoutées, et redistribuées pour maintenir la stabilité numérique pendant l'évolution du contour. Cette stratégie de calcul est rarement intuitive et évidente à développer.

Suivi temporel de structures vasculaires à travers une séquence d'images

Le suivi automatique de structures vasculaires à travers une séquence angiographique présente aussi plusieurs défis dus aux variations d'intensités à travers la séquence. En effet, l'utilisation d'un agent de contraste pour visualiser les artères implique des variations d'intensités dues à l'avancement de l'agent de contraste dans le sang des artères. D'autre part, les mouvements complexes des structures d'intérêts peuvent facilement induire en erreur les algorithmes de suivi basés sur le contenu fluoroscopique des images.

Plusieurs travaux ont été aussi proposés pour le suivi temporel des artères coronaires. Les approches basées sur la mise en correspondance telles que le filtre de Kalman, proposé par Curwen et al. [41], sont probablement les plus applicables dans le contexte de mise en correspondance temporelle d'un ensemble de primitives qui doivent être suivies à travers la séquence. Par contre, la supposition que le niveau de bruit des images soit connu *a priori* rend la technique difficile à implémenter. Dans la classe d'approches de suivi de contours, Meunier et al. [39] utilisent le flux optique; une méthode qui utilise une métrique basée sur l'intensité de l'image pour suivre une région d'intérêt. Une contrainte de lissage sur les champs de vitesse pour les pixels appartenant à la région d'intérêt est utilisée pour améliorer le suivi. Par contre, le flux optique classique se détériore quand les mouvements d'objets ont un très grand déplacement entre les images (>10pixels). Des méthodes par fenêtres de corrélations ont été proposées par Bellemare et al. [49] pour suivre les déplacements des bifurcations d'artères. Cependant, l'intensité ne doit pas varier énormément entre deux images consécutives afin que la fenêtre de corrélation puisse retrouver la position des bifurcations dans une image subséquente.

Calibrage et reconstruction 3D biplan

La méthodologie à suivre pour obtenir une reconstruction tridimensionnelle peut être décomposée en quatre étapes: (i) une calibration permettant de récupérer la transformation pour passer du système de coordonnées d'une caméra à une autre, (ii) la mise en correspondance dans chaque image entre les pixels et les points réels (iii) une

projection inverse pour retrouver le point 3D à partir des deux points 2D et (iv) une interpolation afin de reconstruire les points 3D manquants si la mise en correspondance est éparse. Cependant, la plupart des algorithmes de reconstruction biplanaires s'appuient sur le concept de géométrie épipolaire pour la mise en correspondance des repères dans les images 2D. Pour chaque point de l'image de référence le point correspondant dans l'autre image se trouve sur la droite épipolaire. Lorsque cette droite ne rencontre aucun point, le point le plus proche de la droite épipolaire est choisi comme point de correspondance.

Les travaux de Cheriet et al. [101] ont démontré que la reconstruction 3D biplanaire peut également être réalisée en utilisant une méthode de correspondance temporelle s'appuyant sur cinq paires d'images angiographiques consécutives. Les points reconstruits sont les bifurcations des artères visibles dans les images et le nombre de points nécessaires minimal est de six sur chaque paire. Ainsi une mise en correspondance temporelle des six points de bifurcation à travers la séquence permet d'assurer une stabilité de l'algorithme même si le nombre de correspondances stéréoscopiques est insuffisant. Néanmoins, la possibilité de détecter d'autres repères sur les contours ou l'axe centrale des artères est nécessaire car selon l'angle d'incidence du système à rayons-X, nous ne pouvons garantir la présence de plusieurs bifurcations dans les images.

Reconstruction 3D monoplanaire

Dans le domaine de la vision, la plupart des algorithmes proposés pour la reconstruction d'une scène ou d'un objet prennent en considération que le mouvement réel est strictement rigide. Néanmoins, les concepts de reconstruction pour une séquence monoplanaire restent sensiblement les mêmes que pour l'analyse de deux vues. Nous supposons aussi que les paramètres de la caméra ne changent pas et que les points correspondants entre deux images consécutives peuvent être triés par une méthode appelée RANSAC (random sampling consensus algorithm). En appliquant les concepts de la matrice fondamentale caractérisant l'information contenue dans la scène, nous

pouvons déduire que deux images monoplanaires consécutives peuvent être représentées par deux caméras distinctes. Nous faisons pour cela les hypothèses que l'objet de la scène est resté fixe et que nous avons bougé la camera à deux occasions.

Dans un contexte clinique, Sherknies et al. [117] ont proposé un algorithme de reconstruction monoplanaire d'un transducteur au bout d'un cathéter IVUS (intravascular ultrasound) tenant compte de la vitesse de rétrogradation du cathéter. Ceci permet de connaître correctement le déplacement 3D du transducteur entre deux images consécutives de la séquence d'angiographie. La reconstruction 3D du transducteur peut alors être estimée assez précisément. Finalement, Fallavollita et al. [119] ont développé une méthode d'estimation de la profondeur d'une électrode sur une seule image. Ils calculent la largeur en pixels de l'électrode et utilisent les valeurs de la distance focale et de l'angle du système à rayons-X pour estimer la profondeur.

Les rares méthodes proposées pour la reconstruction monoplanaire où un mouvement non-rigide a été considéré utilisent des hypothèses sur la géométrie de l'objet ou son positionnement à certains instants pour reconstruire correctement l'objet en 3D.

Méthodologie

Lors d'une première étape, l'image angiographique représentant la diastole est extraite en calculant l'écart-type de la différence entre deux images consécutives. L'écart-type minimal correspond au minimum de mouvement possible entre deux images et donc, peut être caractérisé par la phase diastolique du cycle cardiaque.

Deuxièmement, l'image diastolique est filtrée par quatre filtres différents: homomorphique, anisotropique, choc et morphologique. Ces filtres sont utilisés pour la restauration et le rehaussement des images angiographiques. En effet, les images angiographiques sont fortement bruitées et un flou local est souvent présent à cause des artefacts de mouvement et des variations d'intensité dues à l'avancement de l'agent de contraste dans le sang. Ces quatre filtres ont le même effet que le filtre proposé par Frangi. Cependant, pour notre application ces filtres doivent être implémentés sans avoir

à changer leurs paramètres continuellement (filtre automatique). Étant intéressés par la reconstruction de la ligne centrale des vaisseaux ou des cathéters utilisés lors des interventions, nous utilisons une méthode du « Fast Marching Method » . Cette méthode offre à l'utilisateur le choix de la sélection de deux points sur une artère spécifique et d'extraire automatiquement la ligne centrale entre les deux.

Le suivi automatique de la ligne centrale à travers la séquence d'images est effectué en utilisant une procédure basée sur le flot optique multi-échelle et des contours actifs « Gradient Vector Flow ». Le flux optique multi-échelle permet de calculer les mouvements d'artères sachant qu'ils peuvent effectuer de grands déplacements entre deux images consécutives. Ces estimations de déplacements en pixels, sont additionnées aux positions de la ligne centrale dans la première image pour obtenir les coordonnées approximatives de la ligne centrale dans la deuxième image. Ces coordonnées sont utilisées pour paramétrer le contour actif. Ces étapes sont répétées sur toutes les images angiographiques représentant le cycle cardiaque.

Enfin, l'étape de segmentation doit être faite en temps réel afin d'assurer un guidage interactif durant l'intervention. Une fois les lignes centrales des structures d'intérêt segmentées et suivies sur une séquence monoplan d'images angiographiques, une procédure d'auto-calibrage est développée afin de reconstruire ces lignes centrales en 3D. Premièrement, la reconstruction est réalisée en utilisant la méthode biplanaire proposée par Cheriet et al. Nous ajoutons une contrainte sur la courbure 2D/3D des artères pour filtrer les correspondances trouvées par la méthode géométrique épipolaire. Le choix de la procédure d'auto-calibrage monoplanaire est ensuite fortement motivé par le fait suivant: il est techniquement faisable de déterminer les paramètres intrinsèques et extrinsèques d'un système à rayons-X à partir du contenu naturel des images si un modèle 3D *a priori* des artères a un instant donné est disponible. De plus, il a été clairement démontré que le contenu temporel des séquences peut être exploité afin de compléter l'information spatiale. L'auto-calibrage du système d'angiographie est donc basé sur un appariement temporel de primitives géométriques extraites de façon

automatique à partir des séquences monoplan d'images. Les paramètres intrinsèques et extrinsèques du système sont fixés pour cette première analyse.

Résultats et Discussion

Les résultats du filtrage des images angiographiques dépendent fortement des valeurs des divers paramètres associés aux quatre filtres choisis. Nous avons déterminé tous les paramètres empiriquement sur la base de nombreuses expériences. Le rehaussement automatique de l'artère principale, dans les sept images angiographiques utilisées, a été très satisfaisant. De plus, la méthode du FMM a permis d'extraire correctement la ligne centrale entre deux points sélectionnés par l'utilisateur. En calculant la distance entre les coordonnées obtenues par le FMM et la ligne tracée manuellement par un observateur nous avons observé que notre filtre fonctionne aussi bien que celui de Frangin en étant automatique.

La méthode de suivi temporel a été évaluée sur 38 images angiographiques, provenant de trois différents ensembles de données d'images. Premièrement, nous avons remarqué une nette amélioration de l'estimation des déplacements par la méthode du flot optique multi-échelle comparativement à la méthode traditionnelle. Ceci est très avantageux car les coordonnées du contour actif estimées dans les images successives sont proches de la ligne centrale réelle de l'artère coronaire. L'optimisation permet au contour de converger avec succès dans 92% des images. Dans les trois cas négatifs, le contour à convergé vers d'autres structures adjacentes à l'artère choisie.

La reconstruction 3D des artères en utilisant la méthode biplanaire donne des résultats de précision de 1mm à 2mm. Cette précision est directement liée à l'exactitude des correspondances entre les deux vues. En ajoutant une contrainte de triage pour la correspondance, qui spécifie la relation entre la courbure 2D des repères dans les images et la courbure de Frenet-Serret en 3D, les résultats de reconstruction ont été améliorés. En utilisant cinq paires d'images consécutives, pour une configuration antéro-postérieure l'erreur 3D RMS a diminué de 2.8mm à 1.1mm, et pour une configuration

oblique antérieure droite/oblique antérieure gauche l'erreur a diminué de 3.1mm à 1.9mm.

Finalement, nous proposons la stratégie suivante pour la reconstruction 3D monoplanaire des artères (ou cathéter). En supposant que nous disposions du modèle 3D de la structure d'intérêt à un instant donné, il est possible d'estimer la géométrie 3D de la structure en utilisant un minimum de trois images rayons-X consécutives. De plus, la méthodologie devient plus précise en exploitant les sections de structure comportant des mouvements strictement rigides, tel les bifurcations ou électrodes du cathéter. Une étude clinique a été effectuée pour l'estimation de la profondeur d'une électrode sur trois images consécutives. Pour une configuration antéro-postérieure l'erreur 3D moyenne et l'erreur maximale sur la profondeur est de 2.68mm et 7.05mm respectivement, et pour une configuration oblique antérieure droite/oblique antérieure gauche l'erreur moyenne est de 2.34mm et l'erreur maximale est de 5.84mm. De futures études doivent mettre l'emphase sur des contraintes de non-rigidité pour tenir compte de la contractilité épicardique en plus du mouvement global du cœur, des structures anatomiques et des instruments.

Conclusion

La recherche proposée ici établira les bases scientifiques d'outils de visualisation 2D et 3D qui visent à faciliter l'intervention et d'assister les cardiologues dans ces procédures. Les deux contributions principales de ce travail de recherche ont été vérifiées par des expérimentations synthétiques et ont été validées sur des données réelles. Nous pouvons donc conclure que nous avons répondu aux objectifs de recherche établis au début de ce travail.

La première contribution a permis de montrer qu'une contrainte géométrique a permis de raffiner la mise en correspondance de points sur une paire de vues non simulatanées pour améliorer la précision de la reconstruction 3D des artères. Dans le cadre des expérimentations synthétiques, on a ajouté du bruit simulé d'un ordre de 5 pixels d'écart-type sur les coordonnées 2D des structures projetées et les résultats

prouvent que la contrainte géométrique de courbure améliore la précision de la reconstruction des artères. En raffinant les correspondances obtenues par la méthode de RANSAC, nous réduisons les erreurs 3D RMS de 1.61mm à 1.44mm en utilisant la paire de vue postéro-antérieure et latérale et de 2.35mm à 1.86mm en utilisant la paire de vue oblique antérieure droite et gauche.

La deuxième contribution a permi de démontrer qu'il y a possibilité de reconstruire des structures en utilisant seulement une séquence angiographique et un modèle 3D a priori à un instant donné des structures d'intérêt. Cette innovation a démontré que nos équations développées pour récupérer la structure 3D à travers des séquences d'images monoplanaires vise des segments d'artères ou des cathéters qui subissent seulement des mouvements rigides afin d'exploiter la non-contractilité de ces structures. En utilisant un minimum de trois images consécutives jusqu'à un total de six images, notre technique obtient des erreurs 3D RMS qui varient de 1mm a 2.5mm pour des mouvements rigides et de 2.5mm a 5.5mm pour des mouvements non-rigide sur des expérimentations synthétiques. Par contre, la précision obtenue par nos équations monoplanaires se dégradent à 8mm si les approximations initiales des déplacements en 3D à travers la séquence sont irréalistes au début du processus d'optimisation. En revanche, la force de notre algorithme monoplanaire repose sur le fait qu'un modèle a priori 3D des artères peut facilement être calculé a partir de deux rotations du fluoroscope avant que la chirurgie commence. Il suffit ensuite d'exploiter seulement les postions spatiales des artères contenues directement dans les images monoplanaires pour estimer les coordonnées 3D de celle-ci à travers la séquence.

Dans une phase ultérieure le modèle géométrique en 3D obtenu pourrait être transformé en modèle biomécanique en intégrant les propriétés mécaniques des tissus et des instruments utilisés lors des interventions chirurgicales. La planification rationnelle de l'intervention chirurgicale par l'utilisation de techniques de modélisation permettra de décider, à l'avance, de déterminer les paramètres optimaux de cette intervention (propriétés mécaniques d'un « stent » à utiliser dans le cadre d'une intervention d'angioplastie, procédure optimale de reconstruction ou de remplacement d'une aorte

affectée, etc.). Ceci aura pour effet de diminuer le temps opératoire, de convalescence ou de réadaptation par une intervention plus ciblée et un personnel opératoire mieux préparé, de diminuer les risques de complications chirurgicales (pertes sanguines, infections,...) associées à la prolongation du temps opératoire. De même dans le cas d'une intervention d'ablation d'un site d'arythmie cardiaque la diminution du temps opératoire entraîne aussi une diminution du temps d'exposition aux radiations étant donné que le guidage de l'intervention est effectué par l'acquisition d'images à rayons-X tout au long de l'intervention.

En termes de travaux futurs, certaines recommandations importantes sur ce travail peuvent être formulées. D'abord, l'extraction complète des contours d'artères coronaires doit être considérée afin de bien représenter la géométrie de l'artère. Les techniques des « level sets » peuvent être utilisées comme algorithme de départ pour cet objectif. Il serait aussi intéressant d'introduire de l'information complémentaire des contours d'instruments ou d'artères en utilisant des modalités CT ou IRM. En effet, la technique de recalage élastique spatiotemporelle des images IRM du patient, qui peuvent être acquises avant l'intervention, permettra d'obtenir un modèle 4D du coeur avec une résolution temporelle équivalente à la résolution des images fluoroscopiques. De futures études doivent mettre l'emphase sur des contraintes de non-rigidité pour tenir compte du mouvement général du coeur, des structures anatomiques, et des instruments afin que la reconstruction 3D monoplanaire puisse devenir une réalité dans les applications cliniques.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF NOTATIONS

DEFINITIONS

CVD	cardiovascular disease
CT	computerized tomography
MRI	magnetic resonance imaging
SA	sinus node
AV	atrioventricular node
2D	two dimensional
3D	three dimensional
SID	source-to-image distance
SOD	source-to-object distance
PA	primary angle
SA	secondary angle
σ_x	principal point in x
σ_y	principal point in y

SYMBOLS

σ	scale factor for gaussian convolution
γ_H	homomorphic filter high frequency cutoff
γ_L	homomorphic filter low frequency cutoff
I	image
$c(x,y;t)$	anisotropic diffusion coefficient
∇	image gradient
a	complex shock filter slope sharpness
λ	complex shock filter scalar
$\tilde{\lambda}$	complex shock real scalar
ξ	direction perpendicular to gradient
η	direction along gradient

τ	fast marching method cost image
I_x	optical flow spatial gradient in x-direction
I_y	optical flow spatial gradient in y-direction
I_t	optical flow temporal gradient
α	"snakes" 1st derivative parameter
β	"snakes" 2nd derivative parameter
ρ	confidence value for RANSAC
s	# of samples of RANSAC
κ	3D curvature

Chapter 1- Introduction

Cardiovascular disease (CVD) is the single leading cause of death in America today. It accounted for 452,327 of all deaths in the United States in 2004 and 72,338 in Canada that same year [1]. Cardiovascular disease is defined as an injury or lesion within the heart that affects blood vessels, veins and coronary arteries. It is manifested in different forms such as heart attack, angina, or atherosclerosis. Since the risk factors seem to be very much related, it is rather difficult to determine where the damage really begins. One area to focus on is atherosclerosis, which results from the buildup of plaque in the arteries when fatty substances (lipids) gradually accumulate within the inner walls of the arteries. Eventually, the accumulation grows large enough to harden the plaque and almost completely block the blood flow within the artery. The partial or complete blockage within the artery is also known as stenosis that creates drastic complications to a patient's health and leads to one of the possible forms of CVD mentioned above.

Possible remedies for individuals diagnosed with heavy arterial blocking include medication (anticoagulants); and quite frequently balloon or stent procedures are applied in order to re-establish normal blood flow in the blocked artery. Balloons comprise the majority of interventional procedures. These devices are inflated to compress the plaque against the artery wall, in a procedure known as angioplasty. In many other interventions, a stent is also used, usually following a balloon angioplasty. Although the stent is utilized as the primary device, the process still involves a balloon, for the stent itself is mounted on an angioplasty balloon so that it can be delivered and deployed at the focus area. The balloon is inflated, and so is the stent. When the balloon is deflated and withdrawn, the stent remains in place, serving as permanent scaffolding for the newly widened artery. Within a few weeks, the natural lining of the artery, called the endothelium, grows over the metallic surface of the stent [2].

Angioplasty procedures are generally performed in sterile rooms called fluoroscopic laboratories with the use of an X-ray fluoroscope. The patient is lightly

sedated for comfort, but remains awake in order to respond to various instructions from the interventionist. The first task involves the acquisition of a diagnostic image known as a coronary angiogram. The procedure involves a catheter that is inserted through the femoral vein and is brought to the main artery of the heart, the aorta, up and around the opening of either the left or right coronary artery. A contrast agent is released in order to enhance the visualization of the various coronary arteries on the angiographic images. When viewed under the X-ray fluoroscope, in a majority of cases the angiogram will reveal the possible location of the obstruction located in the coronary artery. However, there is a possibility that an under-estimation of the severity of stenosis is obtained since a coronary vessel wall may appear healthy in a particular viewing angle even though it may contain a substantial amount of plaque.

This is the primary reason why a complete three dimensional representation of the coronary artery tree is essential to assist the cardiologist in making a proper diagnostic during interventional procedures. It is to note that other imaging modality technologies exist, such as computerized tomography (CT) and magnetic resonance imaging (MRI); however, these are used primarily for pre-operative or post-operative assessment purposes and are never used during interventional assistance procedures. It is the fluoroscopic X-ray imaging modality that is commonly used in routine clinical procedures in order to guide interventions such as angioplasty or catheter ablation to name a few.

Three dimensional reconstructions of the coronary arteries can be achieved by acquiring two images using either a biplane or a single C-arm X-ray fluoroscope. Biplane systems are seldom deployed in hospitals due to their relatively high cost. Also, if they are available in clinical institutions they are generally used for diagnostic purposes but never used in assisting the cardiologist during angiographic interventions. Regarding the single C-arm fluoroscopes, these are the more typical systems found in hospital intervention rooms and are primarily used to aid the cardiologist in making proper decision but are not used during interventions for the 3D reconstruction as they must be rotated twice in order to acquire the two images. The acquisition of these two

non-simultaneous images causes certain problems when considering performing 3D reconstruction. They more notably include misalignment artifacts which will produce imprecise 3D results. But more importantly, establishing point correspondence between non-simultaneous views is a difficult task. Knowing that point matches between the coronary arteries in the two views is a requirement in performing 3D reconstruction, the following questions are raised:

a) Is there a way to restrict additional error to the resulting reconstruction by making the point correspondence between the images more precise and optimal?

Furthermore, recalling that rotating the C-arm fluoroscope twice or using a biplane fluoroscope is clinically infeasible to assist the cardiologist during the interventions, a second problem is raised:

b) Can 3D reconstruction be performed using only a single X-ray angiographic acquisition from the C-arm fluoroscope in order to assist cardiologists?

The questions raised above describe two important facets when considering treating coronary heart disease. The first facet revolves around the successful diagnosis of stenosis by the cardiologist whereas the second facet focuses on providing a visual aid to the cardiologist in guiding angioplasty procedures. The major contributions of the following research work will intend to answer the previous two questions by focusing on novel methods that will: (i) allow accurate point correspondences between angiographic images in order to arrive at an accurate 3D reconstruction of the coronary arteries, and (ii) provide a 3D reconstruction of the coronary arteries using only a single angiographic view. If this is a possibility then a few advantages would arise in that the overall intervention time would be diminished, since only one image sequence would be acquired concurrently to assist the cardiologist during the intervention. Second, the amount of X-ray dosage from the source towards the patient would be diminished by at least half the amount when compared to the biplane and C-arm two-rotation procedures.

To arrive at a suitable 3D reconstruction using only a single view the implementation of the following steps are required: (a) calibration of the X-ray system in order to determine accurately its geometric gantry parameters, (b) angiographic X-ray image enhancement in order to improve image quality and reduce possible artifacts present in the images, (c) image segmentation to extract coronary artery centerlines for coronary point correspondence, and (d) track the coronary artery centerlines temporally across the single plain acquisition. These four steps become a requirement to improve the correspondence of coronary points for the eventual reconstruction phase using a two view setup and a novel monoplane setup.

1.1 Cardiac Physiology

The heart is made up of four chambers: two upper atria and two lower ventricles, separated by a wall and valves. When a chamber is relaxed, or expanded, it is in *diastole*, whereas a chamber that is in contraction is said to be in *systole*. The heart's electrical rhythms begin as impulses emitted from the sinus node (SA node), the heart's "natural pacemaker." The SA node is a small cluster of specialized cells located in the right atrium, the upper right chamber of the heart. From the sinus node, the electrical impulses travel across a specific route, or pathway, through the atrioventricular node (AV node) and into the lower chambers of the heart (ventricles). Once they reach the ventricles, the impulses serve as a set of instructions, causing the chambers to contract in a routine and consistent manner. The AV node then sends the signals out to the walls of the ventricles (Figure 1.1- left).

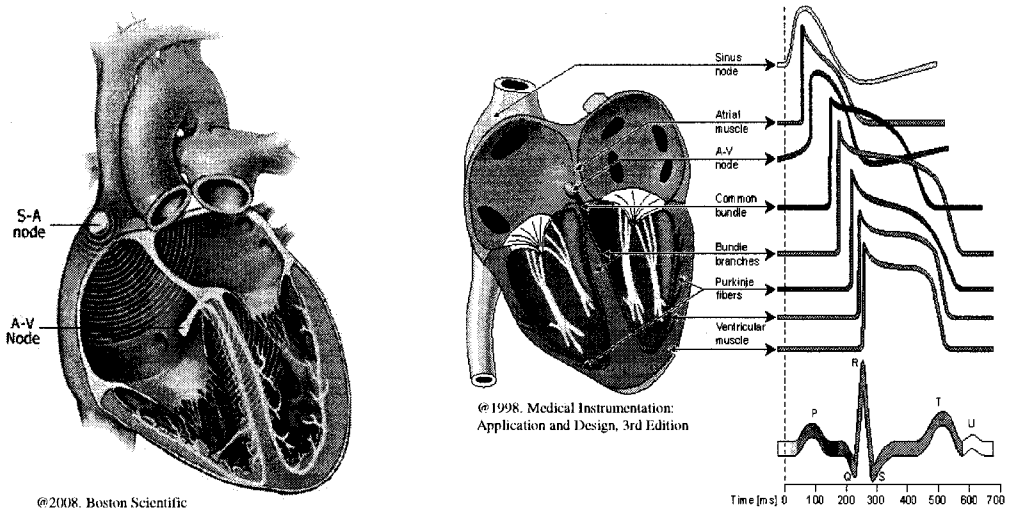


Fig 1.1. (*Left*) Electrical pathway in human heart. (*Right*) The ECG wave composed of the P wave, QRS wave and T wave.

Figure 1.1-right illustrates a normal electrocardiogram (ECG) composed of a P wave, a "QRS complex", and a T wave. At the onset of the heart cycle, impulses from the SA node induce the right atrium to depolarize. This depolarization spreads across the atrial muscle causing atrial contraction, increasing atrial pressure and forcing blood into the ventricles. The ventricular contraction phase of the heart cycle is brought about by depolarization of the ventricles via the atrioventricular (AV) node, Bundle of His and the Purkinje fibers. The AV node provides a delay time allowing the atria to pump blood into the ventricles before ventricular contraction. The Bundle of His and Purkinje fibers allow the ventricles to be depolarized rather instantly. Both the P wave and the components of the QRS complex are depolarization waves. The T wave is caused by currents generated as the ventricles recover from the state of depolarization. This process occurs in the ventricular muscle about 0.25 sec after depolarization and this wave is known as a repolarisation wave. Thus, the electrocardiogram is essentially a representation of depolarization and repolarization waves.

The heartbeats that can be felt at one's pulse rate are formed by the heart's regular rhythmic contractions. When individuals maintain a healthy habit by exercising, eating properly and avoiding tobacco use, they automatically decrease the chances of

developing heart disease. An unhealthy lifestyle or eating habit may lead to a blockage in one or more of the coronary arteries that would require stent or balloon therapy (Figure 1.2).

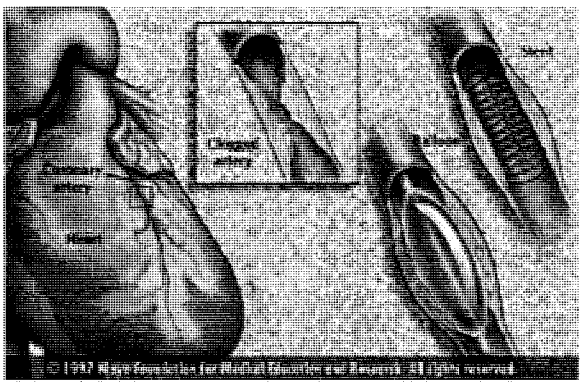


Fig 1.2. (*Left*) The coronary arteries are responsible for directly supplying the heart muscle with fresh blood and nutrients it needs to perform correctly. (*Center*) Coronary artery disease (CAD) begins when hard cholesterol substances (plaques) are deposited within a coronary artery. (*Right*) Catheter balloons are inflated in order to crush the plaque or stents are positioned to re-establish blood flow [3].

1.2 Diagnostic Imaging

In the last 30 years, X-ray angiography has become the method of choice in the diagnosis of coronary heart disease [4]. The 2D angiographic images obtained are either analyzed manually or automatically by algorithms in order to assess the severity of the stenosis, by quantifying the amount that the vessel wall has diminished in width due to the build-up of plaque inside of it (Figure 1.3-left). This implies obtaining an X-ray image from a patient. The X-rays pass through the patient and strike a fluorescent plate (intensifier) that is scanned by a television monitor (Figure 1.3- right). The cardiologist can observe what is happening on the fluoroscopy monitors. One major problem lies with the amount of radiation exposure, both to the patient and to the cardiologist. By minimizing radiation exposure, we in turn diminish the risk of skin injury. In one study, it was estimated that each hour of radiation exposure was associated with an increase in lifetime risk of fatal cancer of 0.1 percent.

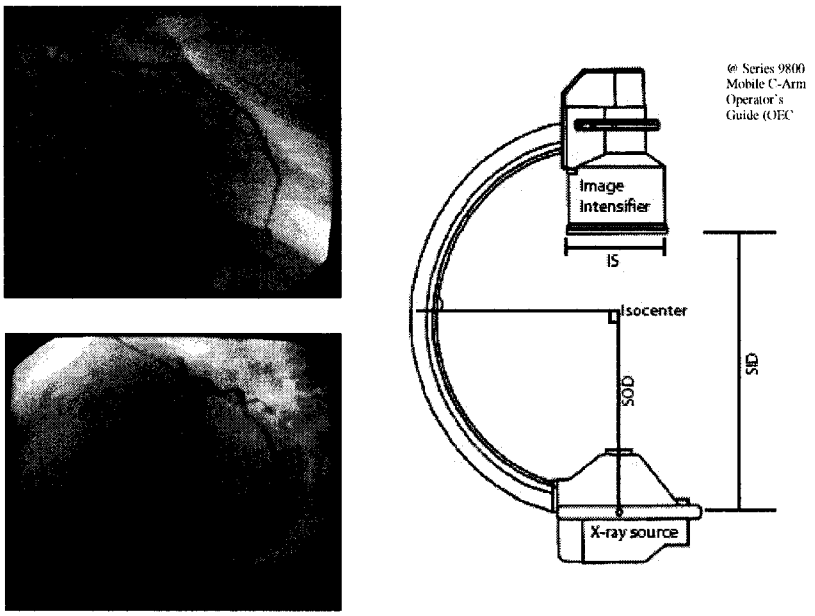


Fig 1.3. (*Left*) Angiographic images of the left and right anterior oblique coronary view. (*Right*) Depiction of a typical C-arm fluoroscope. Some of the gantry parameters are illustrated here, namely the SOD, or source to object distance, the SID is the source to image intensifier distance, and IS is the intensifier size.

Once the heart structure is visible on the monitors a contrast agent is inserted inside the diseased coronary artery so as to enhance its contours. The low signal-to-noise ratios of the 2D angiographic images yields very noisy images that may include artifacts such as motion blur.

1.3 Monoplane X-Ray Fluoroscopy

To guide angioplasty procedures, the cardiologist always uses a monoplane X-ray fluoroscope that is present inside the interventional lab. Depending on the angle of inclination of the C-arm fluoroscope during the procedure, the cardiologist may be looking at an angiographic image that may not fully show a diseased coronary artery. The X-ray dose may pass through the artery in such a way that the severity of the

stenosis may be severely underestimated by the cardiologist. This would result in a false diagnosis.

Another problem arises when guiding the balloon catheter or stent towards the position of the artery blockage. Seeing how imperative it is to position correctly the balloon or stent inside the artery, then how does the cardiologist determine the depth position of the catheter by only making use of the 2D fluoroscopic monitors? Let us take for example the case of another type of intervention that makes use of X-ray fluoroscopy: catheter ablation procedures. A patient that has an arrhythmia will have the electrical activity begin somewhere else inside the heart. A rather severe form of tachyarrhythmia occurs when the heart's electrical signals originate from one of the ventricles, instead of the correct location in the S-A node (Figure 1.4). This causes a type of arrhythmia called ventricular tachycardia (VT) with very fast heartbeats typically between 150 to 200 beats/min.

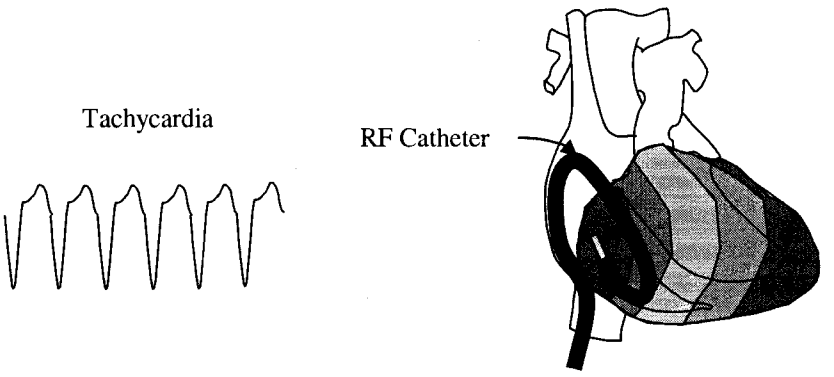


Fig 1.4. Catheter ablation procedures. Patients with sever arrhythmias, such as ventricular tachycardia, have the electrical pathway of the heart begin somewhere inside the ventricles mainly due to diseased heart tissue. An ablation catheter is inserted towards the diseased tissue in order to burn it off and reestablish proper electrical conduction.

Since its initial successful application, RF ablation has become the most common technique for treating cardiac arrhythmias. RF ablation is a technique whereby a physician inserts a catheter through a blood vessel and directs it into a specific location

within the heart. At the tip of the tube is a small electrode, which can deliver radiofrequency energy to burn away the abnormal tissues of the heart that are the root cause of arrhythmia. Once again, the depth of the catheter is of utmost importance for this particular intervention. However, by only making use of the 2D images obtained from the fluoroscope it becomes difficult to determine the exact position of the catheter.

This places importance on providing a 3D visual aid to the cardiologist when performing these types of interventions. In order to reconstruct the ventricle or the diseased coronary artery in three dimensions, a minimum of two views is required. Traditionally, reconstruction is performed using angiographic images that show the diastolic cardiac phase since in this phase there is the least motion experienced by the heart and hence accurate reconstruction will not be hindered by possible artifacts such as motion blur. To verify which angiographic image represents the diastolic cardiac phase an electrocardiogram needs to be acquired concurrently and matched with the images.

1.4 Two x Monoplane X-Ray Fluoroscopy

Most clinical centers have acquired a monoplane C-arm fluoroscope which leads to a sequential acquisition of two non-simultaneous angiographic images synchronized over the electrocardiogram of the patient, in order to use the same two-view reconstruction techniques as the ones applied to the biplane case (Figure 1.5). An example of a C-arm fluoroscope is the *Integris Allura*, from Philips Medical Solutions [5]. This said the precision of the 3D reconstruction obtained is limited due to the different artifacts present during acquisition such as the motion artifacts encountered when turning the C-arm fluoroscope to acquire the second set of images, or the movement artifacts by the patient's respiration. However, a significant problem results in the point correspondence between coronary arteries visible in the two images since these are non simultaneous to one another.

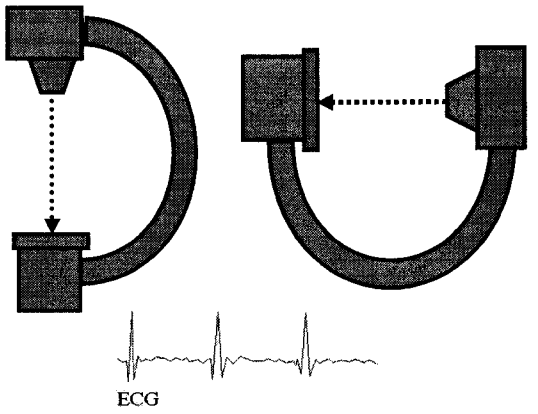


Fig 1.5. Monoplane C-arm fluoroscope. Two angiographic images are acquired at different incidences by rotating the C-arm and synchronizing the images via ECG gating.

1.5 Biplane X-Ray Fluoroscopy

Biplane angiographic systems also exist and have made it possible to reconstruct the coronary arteries in 3D over a few cardiac cycles (Figure 1.6). They allow the acquisition of two quasi-simultaneous images. However the complexity of such biplane systems makes its use in clinical interventions non existent and they are primarily used for clinical research. Lastly, the availability of biplane systems in hospitals is reduced due to the high costs associated with the purchasing of these medical devices.

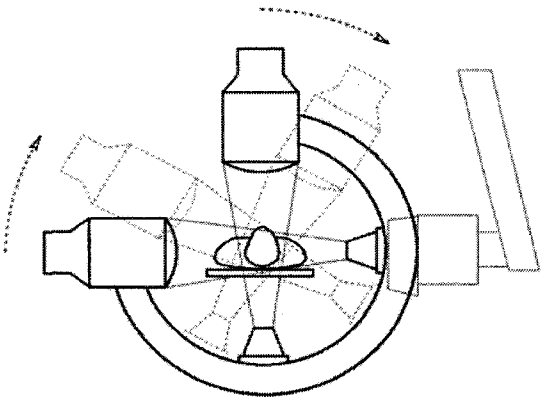


Fig 1.6. Drawing of a biplane X-ray system that provides two perpendicular images quasi-simultaneously at a specific incidence orientation.

1.6 Overall Objective

Vascular imaging has seen many breakthroughs in the detection of coronary disease, in its planning and evaluation for the treatment of this type of pathology. Moreover, the arteries can even be used as minimally invasive conduits for the injection of a specific treatment in the heart. Thus, new interventional applications impose new constraints and methods that need to be developed in order to better treat atherosclerosis by maximizing fluoroscopic image analysis during the clinical procedure. The global objective of the proposed research is to elaborate and validate a method that aims to develop such a clinical tool for the 2D and 3D visualization and interactive manipulation of the coronary arteries, in the angiographic images, for diagnosis and interventional purposes. This potential clinical tool should better equip the cardiologists during their angiographic interventions.

We focus on implementing novel techniques that aim at providing a 3D reconstruction of the coronary arteries for the cardiologist concurrently during interventional procedures. This will produce a visual aid that automatically provides the depth position of the cardiac structure of interest. Thus, we will handle the problem of

point correspondence between two angiographic images acquired by rotating the C-arm fluoroscope twice. Further, seeing how the cardiologist makes use of only a single view to guide interventions, we will propose a novel algorithm to reconstruct the coronary arteries in three dimensions using a single plane approach.

The algorithms developed in our study will provide assistance to cardiologists by also allowing them to load angiographic datasets, select and visualize a healthy or diseased coronary artery in a fluoroscopic image, automatically track a targeted coronary artery during a cardiac cycle, and perform 3D reconstruction in order to allow a visualization during the interventional procedure. It is indispensable to ameliorate the automation, feasibility and reproducibility of all the algorithmic steps proposed in this work in order to provide them to clinical institutions for immediate use. By providing clinical assistance to the cardiologist the methods will subsequently diminish overall interventional time and reduce the risk of interventional complications. Lastly, the methods proposed for angioplasty assistance will immediately benefit other types of interventions such as catheter ablation procedures in treating arrhythmias. Similarly to angiography procedures, reducing the overall operating time will also reduce the exposure to radiation since interventional guidance of the ablation catheters is performed by the constant acquisition of X-ray images throughout the procedure.

1.7 Thesis Outline

In Chapter 2 we will focus on a literary review regarding the following topics: i) cardiac image segmentation, ii) temporal tracking, iii) traditional 3D biplane reconstruction algorithms, and iv) 3D monoplane reconstruction algorithms. A critique will follow at the end of each topic description illustrating the possible disadvantages that each one may have.

In Chapter 3, the specific objectives and research hypothesis will be outlined. Here, we will include the solutions and strategies we propose to remedy the outlined problems.

In Chapter 4 an outline of the methodology required to meet the research objectives will be presented. The first half of the chapter will focus on the tools required to perform 3D reconstruction. Hence, theoretical aspects and equations will be developed for the segmentation, centerline extraction and temporal tracking algorithms we propose. The last part of the chapter will deal with the 3D reconstruction of the coronary arteries using the classical biplane algorithms and a novel monoplane algorithm will be described in detail.

Chapter 5 will be dedicated to the presentation of synthetic and clinical results for the image segmentation, centerline extraction, biplane reconstruction and monoplane reconstruction topics. A discussion will follow each section in order to interpret the findings obtained from our experimental simulations.

Finally, a conclusion is presented that will summarize the various results obtained in our doctoral research and critique their pertinence with respect to the clinical setting. We will discuss the future perspectives regarding our work and their economic and social impact on society, and we will outline possible strategies to tackle the monoplane reconstruction problem in order to make it viable in the near future.

Chapter 2- Review of Literature

This chapter will examine the current literature, while focusing on the contributions of authors in the computer vision field and that of cardiology, more specifically on the geometrical quantification and analysis of coronary arteries. The chapter will be divided into five sections, each related to the overall research objectives outlined in the previous chapter: (i) enhancement of X-ray images, (ii) segmentation of structures in 2D images, (iii) temporal tracking methods, (iv) calibration and biplane reconstruction, and (iv) single view (monoplane) reconstruction.

2.1 Multiscale Enhancement Filters

X-ray fluoroscopy images are characterized by their low signal-to-noise ratio, and therefore image artifacts are certainly present when looking at angiographic images. Therefore, it has become mandatory to improve image quality before performing segmentation. Enhancement filters are used to upgrade image quality as a first step before extracting targeted structures, such as the coronary arteries, within the image.

2.1.1 Lorenz and Frangi Enhancement Filters

In recent years, multi-scale enhancement filtering has been the method of choice for researchers exploring ways to enhance coronary artery contours on fluoroscopic images. Most popular methods are the Hessian-based vessel enhancement filters that have been proposed by Lorenz et al. [6], and Frangi et al. [7]. They both start from the definition of the *scale-space representation L*: $R^2 \times R^+ \rightarrow R$

$$L(x, y; t) = \quad g(x, y; t) * f(x, y) \tag{2.1}$$

where $g(x, y; t)$ is a Gaussian function with variance t, f is an image, (x, y) is a pixel location, and $*$ represents the convolution operation. The Hessian of an intensity image in scale space can be obtained at each point by computing

$$H(x,y)=\begin{bmatrix}\frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x \partial y}\\ \frac{\partial^2 L}{\partial y \partial x} & \frac{\partial^2 L}{\partial y^2}\end{bmatrix}=\begin{bmatrix}L_{xx} & L_{xy}\\ L_{yx} & L_{yy}\end{bmatrix} \tag{2.2}$$

The partial derivatives are computed by convolving the image f with a derivative-of-Gaussian kernel. Frangi proposed computing the ratio of the eigenvalues $|\lambda_1| \leq |\lambda_2|$ of the Hessian, as a ridgeness score $R_B = \lambda_1/\lambda_2$, computing the Frobenius norm, S, of the Hessian to measure overall strength, and combining these into a "vesselness" measure. This measure, specialized to 2D images where vessels are darker than the background, is

$$v_0 = \begin{cases} 0 & \lambda_2 > 0 \\ \exp\left(\left(-\frac{{R_B}^2}{2\beta^2}\right)\left(1-\exp\left(-\frac{S^2}{2c^2}\right)\right)\right) & otherwise \end{cases} \tag{2.3}$$

with parameters $\beta = 0.5$ and c equal to half of the maximum Frobenius norm of the Hessian. The difference between the Lorenz and Frangi vessel enhancement algorithms is that Frangi takes into account all eigenvalues for the response of the filter calculation (Figure 2.1) whereas Lorenz takes into account only the predominant one. This is important as Lorenz's method will only enhance arteries with larger vessel diameters.

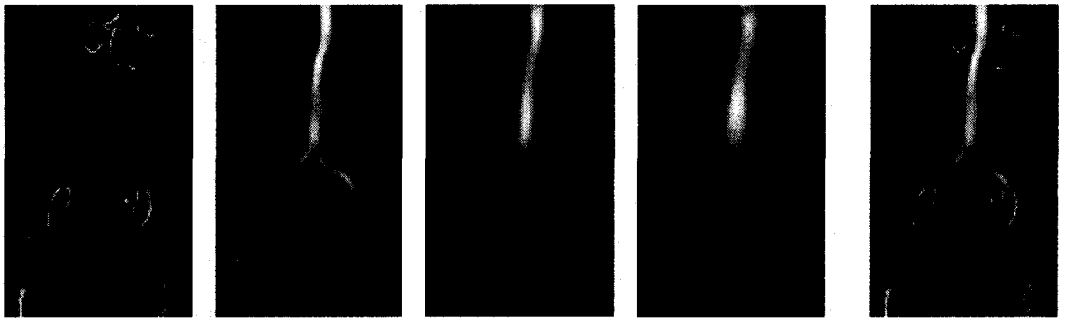


Fig 2.1. The Frangi vessel enhancement filter. The first four images show examples of images obtained by increasing the value of the scale. The last image is the combination of these multiscale measurements by using a scale selection procedure. Image taken from [7].

2.1.2 Vessel Enhancement Limitations

The method of choice in literature for image enhancement is the vessel enhancement filter of Frangi et al. The advantage of the filter is that tubular structures such as coronaries will be enhanced with respect to the background of the image. The mathematical formulation is also simple to implement. However, a major disadvantage of the filter is that the scale σ, needs to be chosen adequately in order to enhance the artery of choice. Inadequate selection of the scale values will lead to missing contour information for smaller and larger arteries in the angiographic images. The selection of the scale factors depends on the number of larger or smaller vessels present in the images. A larger scale factor will influence a positive response to the larger sized arteries. Usually the algorithm will run for various values of the scale (i.e. scales between 0.5-2 pixels) and the final output will reflect the maximum response of the scales used at a given pixel location. Ideally, a segmentation tool should be provided to the interventionist in which an angiographic image is loaded and the major coronary arteries or arteries having a stenosis would be automatically enhanced without the selection of scale factors.

2.2 Segmentation of Angiographic Images

We survey current image segmentation methods, covering both early and recent literature related to vessel segmentation algorithms and techniques. We will elaborate on the following six subcategories: (i) skeleton-based approaches, (ii) region growing approaches, (iii) ridge-based approaches, (iv) mathematical morphology techniques, (v) geometric deformable models and (vi) active contour models. We will end by highlighting a few limitations for this particular procedure.

2.2.1 Skeleton-Based Approaches

These methods target the extraction of vessel centerlines. By connecting together every single extracted centerline we can arrive at a complete two dimensional vessel

representation. This is important as the final vessel coordinates will be used for an eventual three dimensional reconstruction. We first present the 2D skeleton-based approaches as these are the ones that interest us as we are dealing with two dimensional X-ray images, however, we will also briefly highlight relevant techniques applied to the 3D case.

An automatic 2D algorithm is proposed for centerline extraction of coronary arterial tree in X-ray angiographic images. First of all, a Hessian enhancement method is applied, followed by thinning the extracted coronary tree to obtain the skeletons. By applying a centerline correction algorithm, based on Gabor filters, the authors manage to extract both the centerline and diameter of the artery [8]. A multi-features measure that is based on a probability tracking model to extract the blood vessel tree is implemented in [9]. A centerline operator to optimize the vessel skeleton-line is applied to extract the vessel coordinates. Meunier and Bellmore [10] assumed that a coronary bifurcation can be represented by a simple Y geometric structure. Therefore a Fuzzy C-means algorithm is first used to segment the coronary bifurcations. Then the segmented bifurcation is skeletonized to produce the expected Y shape geometry. To define the Y shape geometrically, its center and the branch angles are computed. An algorithm for the automated extraction of the 2D skeletons and borders of coronary arteries in digitized angiograms is proposed by [11]. Initially, the approximate skeleton and borders of the coronary artery tree are extracted through a template analysis. The skeleton and borders of each artery segment are then used for constructing its enclosing area where the defined skeleton and border curves are considered as markers. These markers are used to construct a gradient image where all pixels inside the enclosed area, except skeleton ones, are assigned the gradient magnitude of the original image. An optimization method is applied for extracting the artery segment borders. Lastly, [12] describe techniques to determine the skeleton of coronary arteries in cineangiograms. The technique starts with the detection of skeleton points from a local ridge detector using density profiles of the arteries. Local orientations of skeleton points are obtained directly from the ridge detection or from local edge directions. Boundaries of the arteries are computed from the

density profile along a straight line perpendicular to the skeleton. From this, a sequential and raster-scan tracking procedure is applied to obtain segments of the arteries from two scanning directions. The final 2D vessel skeleton is generated by merging all the segments.

Regarding the 3D skeleton-based methods, Tozaki et al. extract bronchus and blood vessels from thin slice CT images of the lung for 3D visualization and analysis. First, a threshold is used to segment the images. Then, blood vessels and bronchus are differentiated by using their anatomical character. Finally, a 3D thinning algorithm is applied to extract the vessel centerlines. The resulting centerline structure is used to analyze and classify the blood vessels [13]. Kawata et al. analyze blood vessel structures and detect blood vessel diseases from cone-beam CT images. X-ray digital angiograms are collected using rotational angiography. 3D image reconstruction is performed by a short scan cone-beam filtered backprojection algorithm based on the short injection time of the contrast medium. First, a graph description procedure extracts the curvilinear centerline structures of the vessel tree using thresholding, elimination of the small connected components, and 3D fusion processes. Then, a 3D surface representation procedure extracts the characteristics of convex and concave shapes on blood vessel surface [14]. Niki et al. describe a 3D blood vessel reconstruction and analysis method. Vessel reconstruction is achieved on short scan cone-beam filtered back-propagation reconstruction algorithm. A 3D thresholding and 3D object connectivity procedure are applied to the resulting reconstructed images for the visualization and analysis process. A 3D graph description of blood vessels is used to represent the vessel anatomical structure [15]. Sorantin et al. uses a 3D skeleton method in the assessment of tracheal stenosis on spiral CT images. The system extracts the trachea as a single object starting from a user specified seed point. A 3D dilation is employed to handle the uncertain boundary points due to partial volume effect. Via linear interpolation, the 3D volume is converted into cubic voxels. Lastly, the centerline axis is separated from the extracted skeleton using a shortest path-searching algorithm. This step requires the user to mark beginning and end points on the central path [16].

2.2.2 Ridge Based Approaches

These methods will treat a grayscale image as a three dimensional elevation map in which intensity ridges approximate the skeleton of the tubular object in question. After creating the intensity map, ridge points are targeted as local peaks in the direction of maximal surface gradient, and can be obtained by tracing the intensity map from an arbitrary point, along the steepest ascent direction.

Aylward et al. approximate the center axes of tubular objects such as vessels in an angiogram as "intensity ridges". Ridges are tracked by estimating the local vessel directions. First, image intensity is mapped to height to create intensity height surface. Second, from a user-supplied starting point an initial ridge point is found using a conjugate directions search with respect to the Hessian matrix. Third, the ridge is tracked. Finally, the vessel widths are estimated using points on the ridges [17]. The authors show results of a vascular tree extracted from a MR angiogram in Figure 2.2. This required a fair amount of user intervention (105 mouse clicks in all). The method was completed by Bullitt and Aylward by having the final segmented 3D vessel tree be represented by a graph where each individual vessel structure retains information about its relationship to its neighboring vessels [18].

Guo and Richardson propose a ridge extraction method that treats digitized angiograms as height maps and the centerlines of the vessels as ridges. The image is preprocessed using a median filter and then smoothed by an anisotropic diffusion method. Second, a region of interest is selected by adaptive thresholding. This process reduces false ridges introduced by noise. Third, the ridge detection process is applied to extract the vessel centerlines. Finally, the candidate vessel centerlines are connected using a curve relaxation process [19].

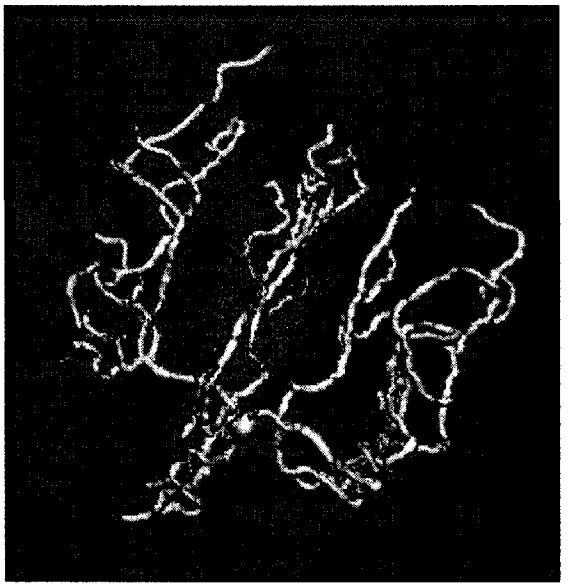


Fig 2.2. The tubular vessels are extracted using Aylward et al. algorithm from 105 user defined clicks. Image taken from [17].

2.2.3 Region Growing Approaches

Starting from some user defined seed points, region growing techniques segment images by incrementally recruiting pixels to a region based on some predefined criteria. Two of the more important ones are value similarity and spatial proximity. Both assume that 2D pixels that lie close to one another and that display similar intensity values are likely to belong to the same vessel. The main disadvantage of region growing approach is that it often requires many user-supplied seed points. Further, due to the variations in image intensities and noise, region growing can result in holes and over-segmentation. Thus, it requires post-processing of the segmentation result.

Schmitt et al. determine contrast agent propagation in 3D rotational X-ray angiography (XRA) image volumes. They combine thresholding with a region growing technique to segment vessel tree in 3D. After the segmentation, propagation information is mapped from the 2D projections to the 3D image data [20].

O'Brien and Ezquerra develop a method to automatically segment coronary vessels in angiograms based on spatial and structural constraints. Like most algorithms a preprocessing step is applied to the image using a low pass filtering. The initial segmentation process begins with the user selecting a single seed point. A region growing process is initialized to extract the initial approximation of the vessel structure from the selected seed point. Using this approximation, a skeleton process starts to extract the centerlines of the structure by employing the following test: a disk is expanded outward until a pixel is encountered which is not part of the extracted region. This case is called a hit. If a hit occurs on opposite sides of the structure having the same distance, then this point is selected as a centerline. Lastly, undetected vessel segments are located by a spatial expansion algorithm. At this stage, images are divided into two categories: background and vessel structures [21]. Figure 2.3 shows the result of their method applied to an angiogram image.

Fig 2.3. (*Left*) The original angiographic image. (*Center*) Initial approximation of the vessel, and (*Right*) the superimposed segmentation using O'Brien and Ezquerra's method. Image taken from [21].

Yim et al. present a grayscale skeleton method for the segmentation of vessel tree structures using magnetic resonance angiography (MRA). Their method is based on the ordered region growing (ORG) algorithm which uses the connectivity between all the voxels in the image and transforms it into a graph. A distinctive feature that separates this method from other graph-based methods is that the path used has minimal dependence on initial seed location. Here, the skeleton process is performed in two

ways. First, a user explicitly selects the origin of the vessel tree and endpoints of every vessel. Then, vessel segments are extracted by tracing the path from each endpoint to the origin of the graph. The second method is a pruning process based on the branch length to discard false vessel branches [22].

2.2.4 Morphological Techniques

Morphology relates to the study of object forms or shapes. Morphological operators apply structuring elements to images. These elements are matrices containing ones and zeros and may have different shapes such as disks, rectangular, or diamond to name a few. Dilation and erosion are the two main operators. Dilation expands objects if they lie inside the defined structuring element, filling holes and connecting disjoint regions. Erosion shrinks objects if they lie outside the structuring element.

Eiho and Qian propose a method based on pure morphological operators for the detection of the entire coronary artery tree visible in two dimensional cineangiograms. First, a "Top-hat operator", which is able to detect the local elevations on arbitrary backgrounds, is applied to enhance the shape of the vessels. Second, morphological erosion followed by half-thresholding operations are applied to remove the background areas. A user then selects a point anywhere on the artery tree. Following that, the system extracts the entire artery tree using neighbor checking according to the average grayscale intensities between them. Finally, the edges are extracted by applying a watershed transformation on the binary image obtained from a dilation operation [23].

2.2.5 Geometric & Parametric Deformable Models

Deformable methods apply specific vessel derived models to extract either the vessel medial axis or contour perimeter. There are two categories of deformable models: (i) geometric deformable models or front propagation methods, and (ii) parametric deformable models, also known as active contours.

2.2.5.1 Geometric Models

Caselles et al. [24] and Malladi et al. [25] use propagating interfaces under a curvature dependent speed function to model anatomical shapes. They used the Level Set Method approach developed by Osher et al. [26] and adapted it for shape recognition processes. The main idea behind the Level Set Method is to represent propagating curves as the zero level set of a higher dimensional function which is given in the Eulerian coordinate system. Hence, a moving front is captured implicitly by the level set function. It can handle complex interfaces which develop sharp corners and change its topology during the development. Since the level set function is given in the Eulerian coordinate system, discrete grids can be used together with finite differences methods to obtain a numerical approximation to the solution. The propagation of the front through a vessel in an angiogram image is shown in Figure 2.4.

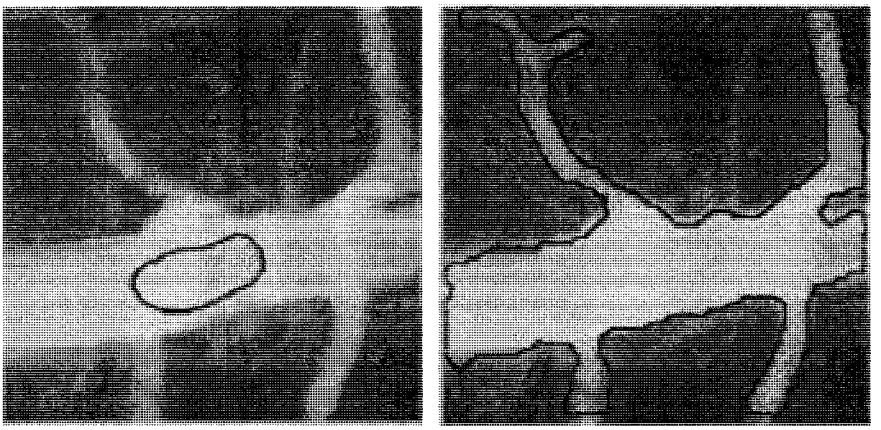


Fig 2.4. Propagation interfaces under a curvature dependent speed function, as described by Malladi et al., and applied to an angiographic image. Image taken from [27].

Sethian also developed another noteworthy technique, called the Fast Marching Method [27], which uses a wave propagation approach for specialized front problems. Fast Marching Methods are used in the problems where the front advances

monotonically with a speed that does not change its sign. The Fast Marching Method's advantage over the Level Set Methods is that it is more computationally efficient.

Several methods have been described [28–29] for the segmentation of the central axis of an artery. In terms of the Fast Marching Method (FMM) approaches, defining a speed function is crucial in path extraction, since the extracted path is a minimum-cost path of the specific cost selected. For a non-trivial path extraction problem, such a cost is often difficult to define. In [29], the cost image was defined as the reciprocal of the vessel-enhanced image. When extracting guidewires or line-like objects, enhancing the X-ray image using non-maximum suppression and a salience distance transform produced a successful cost image [30]. However, some discrepancies between the segmentation and the real vessel tree still persist. A major difficulty is that vessels can be broken into several disconnected components and that discontinuities may occur at bifurcation or stenosis points.

2.2.5.2 Active Contour Models

These models are model based techniques employed for finding object contours or center axis using parametric curves that deform under the influence of internal and external forces. First introduced by Terzopoulos et al., active contour models or *snakes* are a special case of a more general technique of matching a deformable model by means of energy minimization [31]. Physically, a snake is a set of control points, called *snaxels*, in an image that are connected to each other. Each snaxel has an energy associated with it. This energy either rises or falls depending upon the forces that act on that snaxel. These forces are known as snake's *internal* and *external* forces, respectively.

Historically, all active contour models proposed today in 3D vision or medical image analysis have emanated from the classical snakes algorithm. The snake is represented by a curve, $v(s)$, which usually requires user interaction to select many control points defining the initial curve. In this case, the 1^{st} derivative term of the parametric equation is related to the *bendedness* of the snake, whereas, the second

derivative of the curve relates to the *rigidity* of the snake. The classical Snakes equation is as follows:

$$Energy = \int_0^1 E(v(s))ds = \int_0^1 [E_{\text{int}}(v(s)) + E_{ext}(v(s))]ds \tag{2.4}$$

where the internal energy is defined as:

$$E_{\text{int}} = \left(\alpha|v_s(s)|^2 + \beta|v_{ss}(s)|^2\right)/2 \tag{2.5}$$

The internal energy reflects the bending and stiffness of the parametric curve. The external energy always comes from some specific image feature (external force) and it is defined as:

$$E_{ext} = \gamma * Intensity \tag{2.6}$$

That is, the force will be proportional to the feature image pixel intensities and the snake will move from a higher to a lower intensity field. By varying the parameters α, β, γ, the snake can eventually converge on a given boundary or landmark. Another way of defining the internal energy is by taking into account the total length of the initial curve. The internal energy will be defined as a curvature:

$$K = \int_0^1 \frac{|v'(s) \times v''(s)|}{L^3} ds \tag{2.7}$$

where the variable L denotes the length of the curve

$$L = |v'(s)| \tag{2.8}$$

This energy term takes into account the total length of the snake curve. This is important as it allows the user to fine tune the parameters α, β, γ such that the final optimized curve will be roughly the same length as it was prior to the optimization step.

Molina et al. use 3D snakes to reconstruct 3D catheter paths from biplane angiograms. The 3D snake used in this method is represented by B-splines and is initialized interactively. Using a snake facilitates the merging information from both projections simultaneously during the energy minimization process [32].

Klein et al. describe an approach to extract vessels from X-ray angiograms using deformable snakes. The user provides an initial estimate of the location of the vascular entity, and the algorithm refines the estimate by deforming a snake to minimize some energy function. A B-spline model is used in their snake implementation. They use a Gabor filter to determine the image (or edge) energy term to attract the snake. The approach is most suitable for the accurate extraction of vascular segments. However, the amount of user interaction and computation required makes it impractical for extracting entire vascular structures [33].

2.2.6 Segmentation Limitations

Tubular cores requires a fair amount of intervention by selection numerous seed points to activate the algorithm. Morphological operators coupled with binary filtering have strict disadvantages in that the image background will be suppressed with the detriment of losing artery contour information if a proper threshold is not selected adequately. The level set methods have become increasingly popular to extract the entire contour of anatomical structures; however most methods have propagating wave fronts that experience "leakage" near artery perimeters if a regularization term is not chosen appropriately.

Various image details must be taken into account for successful implementation of deformable model segmentation. These details include: (i) the amount of noise present in the image as well as visible external structures such as catheters or spinal cord that may attract the curve during the convergence process, (ii) the complexity of the

defined parametric curve in terms of the number of control points defining it and the amount of parameters that need to be optimized in order to minimize a given energy function, and (iii) whether the capture range (the distance between coronary arteries between two consecutive frames) is significant enough to hinder the snake from converging correctly to the exact coronary location at a subsequent time instant. Primitive active contour methods required user initialization of many control points to define the deforming parametric curve. In the case where the curve did not converge well in a specific image, the user would reinitialize the deformable model by reselecting the control points. This methodology is tedious and inefficient in a clinical context.

2.3 Temporal Tracking Techniques

We highlight relevant tracking concepts in both the computer vision and the clinical field in order to provide the reader with a broader perspective on the limitations that one has with respect to the other.

2.3.1 Computer Vision Tracking Methods

Object tracking is an important task within the field of computer vision. In its simplest form, tracking can be defined as the problem of estimating the trajectory of an object in the image plane as it moves around a scene. One can simplify tracking by imposing constraints on the motion and/or appearance of objects. For example, almost all tracking algorithms assume that the object motion is smooth with no abrupt changes. One can further constrain the object motion to be of constant velocity or constant acceleration based on a priori information. The aim of an object tracker is to generate the trajectory of an object over time by locating its position in every frame of the video [34].

The main tracking categories: point tracking, kernel tracking and silhouette tracking. Objects detected in consecutive frames are represented by points, and the association of the points is based on the previous object state which can include object position and motion. This approach requires an external mechanism to detect the objects

in every frame. An example of such a tracker is the Kalman filter [35]. Kernel refers to the object shape and appearance. For example, the kernel can be a rectangular template or an elliptical shape with an associated histogram. Objects are tracked by computing the motion of the kernel in consecutive frames. The Kanade-Lucas tracker by [36], is an example of a template kernel approach. Silhouette tracking methods use the information encoded inside the object region. This information can be in the form of appearance density and shape models which are usually in the form of edge maps. Given the object models, silhouettes are tracked by either shape matching or contour evolution. Some examples include the Hausdorff distance [37] or the Hough transforms [38].

Significant progress has been made in object tracking during the last few years. Several robust trackers have been developed which can track objects in real time in simple scenarios. However, it is clear that many assumptions are used to make the tracking problem tractable, for example, smoothness of motion, minimal amount of occlusion, illumination constancy, high contrast with respect to background, etc., are violated in many realistic computer vision scenarios and therefore limit a tracker's usefulness in applications like automated surveillance, human computer interaction, video retrieval, traffic monitoring, vehicle navigation, or even clinical applications [34]. The clear distinction between the computer vision applications with respect to the clinical applications is that the latter involves structures that undergo primarily non-rigid movement, and thus, some of the above assumptions must be applied to deal with the problem of tracking the coronary arteries.

2.3.2 Clinical Tracking Methods

The classical method for recovering the motion of the coronaries is to track the motion of the arterial centerlines in the two dimensional (2D) projection images. Optical flow [39], binary image elastic registration [40], Kalman snakes [41], and local space search and graph minimization techniques [42] are some of the methods that have been proposed for 2D vessel tracking. However, tracking the vessels in the projection image space has significant limitations. Multiple vessel overlap is common due to the

projective nature of the imaging modality. Correctly tracking arteries through these regions is a difficult problem, which can only be solved with additional knowledge and regularizing constraints.

The authors in [43] developed an epipolar tracking method that tracked the arteries in six consecutive 2D X-ray images. The proposed method involves three major steps. First, vessel centerlines are detected using segmentation and each vessel is labeled. Second, the epipolar lines for the leading edges in each vessel in the RA0 (or LAO) view image are drawn on the LAO (or RAO) view image, and the distance between the epipolar lines and the leading edges of vessels are measured. In this case, six frames are selected for searching for vessel correspondence. Lastly, the measured distances in six frames are summed and tabulated. The vessel correspondence with minimum total distance is most suitable. The authors in [44, 45] describe a knowledge based system that interprets three dimensional coronary artery movement, using data from digital subtraction 2D angiography image sequences. Dynamic information obtained from artery centerline 3D reconstruction and optical flow estimation, is classified according to experimental evidence indicating that artery displacements are quasi homogeneous by a segment analysis. These facts are then related and interpreted using anatomical functional knowledge provided by a specialist, as well as spatial and temporal knowledge, applying spatio temporal reasoning schemes. Applied knowledge is mainly geometrical, kinetic, temporal and anatomic. Displacement data is analyzed by means of a carefully selected group of dynamic features, which are intended to detect significant events, related to anatomical regions through time. These events are represented as homogeneous segments that provide a detailed description of the arteries centerlines dynamic behavior. Such a representation is original in the sense that identified dynamic features can be, at a reasoning level, related in multiple ways, and the amount of features is not limited by the mechanical model that associates the detected observations. Numerical information contained in the homogeneous segment representation is transformed into symbolic labels, which allow an efficient knowledge-based interpretation using rule-based queries.

Other vessel tracking approaches, starting from an initial point, detect vessel centerlines or contours by analyzing the pixels orthogonal to the tracking direction. Different methods are employed in determining vessel contours or centerlines. Edge detection operation followed by sequential tracing incorporating connectivity information is a straightforward approach. Aylward et al. [17, 18] utilize intensity ridges to approximate the medial axes of tubular objects such as vessels. Some applications achieve sequential contour tracing by incorporating the features, the searching direction, and the search range, detected from previous step into the next step [46]. The initial values of the features are supplied by the user. Fuzzy clustering is another approach to identify vessel segments [47]. After the initial segmentation of the pixels into different regions, a fuzzy tracking algorithm is applied to each candidate vessel region. False candidate vessels are rejected by the algorithm within two or three iterations. Some methods utilize a model in the tracking process and incrementally segment the vessels. A more sophisticated approach on vessel tracking is the use of graph representation [48]. The segmentation process is, then, reduced to finding the optimum path in a graph representation of the image.

Template matching tries to recognize a structural model (template) in an image. The matching method uses the template as a context, which is a priori model. In a 2D framework, the authors in [49] provided a template matching technique to track the bifurcation points along consecutive images (Figure 2.5a) whereas the authors in [50] presented a method combining active contours and image template matching techniques for tracking the coronary arteries. The active contour is modeled as a string of templates and is no longer a simple curve. In [51], a method is presented to extract and track the position of a guide wire during endovascular interventions. A rough estimate of the displacement of the guidewires is obtained using a template matching procedure and the position of the guide wire is determined by fitting a B-spline to a feature image using 4-5 control points. In these methods, large deformations may hinder the reliability of the template position estimates. The calculation of velocity fields was performed using correlation in a small region of interest in [52] to track the targeted cardiac structure. To

improve tracking speed, the authors in [53] extended the formulation of active contours by including an additional force, derived from the optical flow field. In both studies, the traditional optical flow methods (block correlation by Horn and Schunk) worked well since the overall inter-frame movement of the cardiac surface was small.

Fig 2.5a. Bifurcation tracking using template approach. Correlation templates are solved for in order to estimate subsequent bifurcation positions. The right image shows a correlation box that misrepresents a bifurcation position due to the sensitive nature to image noise when considering correlation methods. Image taken from [26].

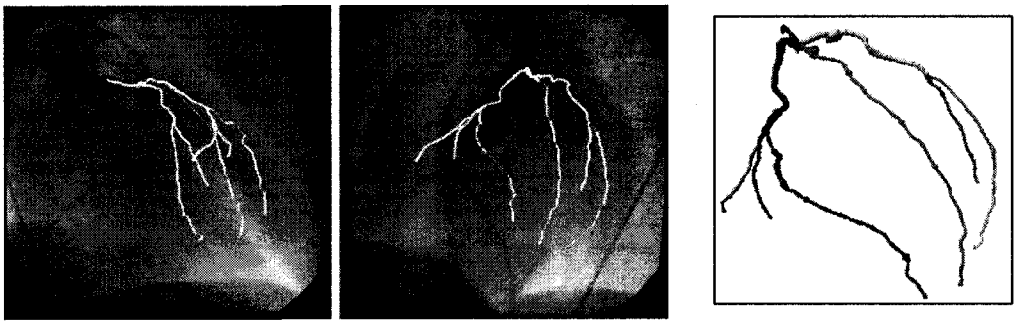


Fig 2.5b. 3D Tracking across biplane images. A 3D model is extracted at a first instant and is modeled using B-Splines. Transformation matrices are registered between the 3D model and the biplane images. Subsequent 3D reconstructions are obtained by solving a coarse to fine energy model which takes into account affine, rigid and non rigid transformation matrices. Image taken from [56].

Making the transition to 3D tracking algorithms, in [54], the authors propose combining optical flow derived displacement information from two projection images to displace an

existing a priori 3D coronary tree model. A deformable model approach for segmenting objects using parametric curves was applied to segmenting the coronary arteries with 3D B-spline curves in [55]. In [56], a registration framework recovers the motion one angiogram at a time, using a set of coarse-to-fine motion models. This 3D approach requires an a priori 3D model of the coronary tree at the first time instant before tracking begins (Figure 2.5b). The main objective of these methods is to evaluate the functionality of the overall coronary motion in three dimensions.

2.3.3 Temporal Tracking Limitations

Template matching techniques and conventional optical flow methods fail when the intensity variations between images fluctuate enormously or when the overall displacements of the object of interest between consecutive images are significant (larger than 10 pixels). Three dimensional tracking methods are more robust than the two dimensional case since an initial 3D model of the arteries is obtained from MRI or CT images. However, these models require many 3D control point initializations and the complexity of the algorithms increase due to the extrapolation errors from local deformations. The mean average time for one time frame was about 155 minutes in [56]. The ideal tracking algorithm will make use of only 2D angiographic sequences, since this is the principal modality for diagnosis, and will limit user interaction for the definition of the deformable model. Ideally it will automatically track the artery in consecutive images regardless of the magnitude of interframe artery displacement.

2.4 Biplane 3D Reconstruction

2.4.1 Biplane X-ray Fluoroscopy Reconstruction Overview

In biplane angiography two prospective projection views of the coronary arteries are acquired using a biplane fluoroscope. The two X-ray sources in a biplane system are related to each other through a rigid body transformation. It can be defined either by the transformation of each single-plane X-ray source relative to a fixed world coordinate

system [57], or by the transformation relating the two single-plane sources [58, 59]. The biplane apparatus consists of two X-ray sources and image intensifiers, at opposite ends of the gantry. Magnification depends on the source-to-object distance (SOD), source-to-intensifier distance (SID), and the intensifier size (IS). In clinical procedures, the patient would be considered the object in this case and so his/her distance from the X-ray source is termed SOD. The SID is also termed as the focal length. Moreover, both X-ray sources rotate around a common isocenter. Images are acquired on the two systems nearly simultaneously (Δt ~10ms). The rotations are described by two angles, a primary (PA) and secondary angle (SA). These inclinations are respect to the patient's location on the intervention table. The PA defines a left-right rotation with respect to the patient whereas the SA defines a cranial-caudal rotation.

To determine the 3D geometry of the coronary arteries using the two perpendicular images following steps are taken: (i) the imaging geometry is determined and calibrated for, (ii) corresponding points in the two systems are determined (iii) vessel profiles are extracted along epipolar lines and then reconstructed. The imaging geometry can be determined using calibration objects or by self-calibrating, that is, optimizing the X-ray parameters using only 2D points contained in both images. Once the geometry is determined, the corresponding points and regions along the coronary arteries are determined. In general, the correspondence is not apparent, such as with bifurcation points; therefore, the correspondence is usually determined using epipolar lines. These concepts will be defined in more detail in the upcoming sections.

2.4.2 Offline and Online Calibration Methods

2.4.2.1 Computer Vision Community

Camera calibration is a necessary step in 3D computer vision in order to extract metric information from 2D images. According to the dimension of the calibration objects, we can classify those techniques roughly into three categories: (i) 3D calibration objects, (ii) 2D plane based calibration, and (iii) self-calibration.

3D Calibration Objects: Camera calibration is performed by observing a calibration object whose geometry in 3D space is known with very good precision. Calibration can be done very efficiently [60]. The calibration object usually consists of two or three planes orthogonal to each other. Sometimes, a plane undergoing a precisely known translation is also used [61], which equivalently provides 3D reference points. This approach requires an expensive calibration apparatus and an elaborate setup. Some existing techniques are Faugeras et al. [62], Weng et al. [63], Wilson [64], and Lenz et al. [65].

2D plane based calibration: Techniques in this category requires observing a planar pattern shown at a few different orientations [66, 67]. Different from Tsai's technique [68, 69], the knowledge of the plane motion is not necessary. Because almost anyone can make such a calibration pattern by him/her-self, the setup is easier for camera calibration.

Self-calibration: Techniques in this category do not use any calibration object, and can be considered as 0D approach because only image point correspondences are required. If images are taken by the same camera with fixed internal parameters, correspondences between three images are sufficient to recover both the internal and external parameters which allow us to reconstruct 3-D structure up to a similarity [70]. Although no calibration objects are necessary, a large number of parameters need to be estimated, resulting in a much harder mathematical problem. Some researchers in the computer vision field proposed specific self-calibration algorithms for restricted motions. In several cases it turns out that simpler algorithms can be obtained. However, the price to pay is that the ambiguity can often not be restricted to metric. Some interesting approaches were proposed by Moons et al. [71] for pure translation, Hartley [72] for pure rotations and by Armstrong et al. [73] for planar motion. Recently some methods were proposed to combine self-calibration with scene constraints. A specific

combination was proposed in [74] to resolve a case with minimal information. Bondyfalat et al. [75] proposed a method of elimination to impose the scene constraints. Liebowitz et al. [76] on the other hand formulate both the scene constraints and the self-calibration constraints as constraints on the absolute conic so that a combined approach is achieved. Another important aspect of the self-calibration problem is that of critical motion sequences. In some cases the motion of the camera is not general enough to allow for self-calibration and an ambiguity remains regarding reconstruction. A first complete analysis for constant camera parameters was given by Sturm [77].

2.4.2.2 Clinical Community

Similar to the computer vision community, an offline and online approach are proposed in order to calibrate the biplane fluoroscopy systems. In the off-line approach, the calibration parameters are computed for a fixed set of C-arm orientations before the surgery begins. The off-line approach allows for larger calibration phantoms, denser grids, and produces images which are simpler to analyze, since only fiducial markers are present in the image. In the on-line approach, the C-arm gantry parameters are computed anew using the angiographic images acquired concurrently during surgery and require a more sophisticated image processing algorithm. It presents trade-offs between calibration phantom size, grid density, and accessibility on the one hand, and robustness and accuracy on the other.

An example of a 3D calibration object used in the clinical community is shown in Figure 2.6. The calibration object usually consists of two or three planes orthogonal to each other and its 3D geometry and lengths are known in order to obtain a precise scale factor when dealing with image pixels and world coordinates.

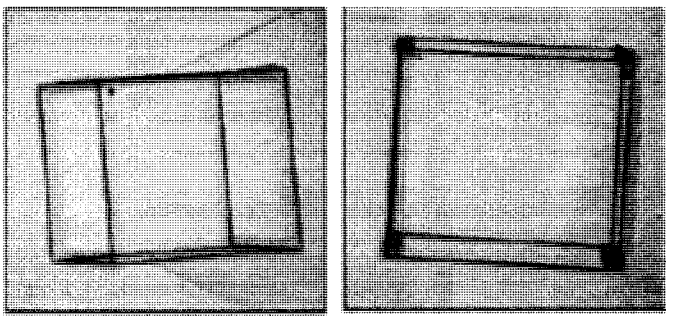


Fig 2.6. Drawing of a typical biplane image calibration cube used for calibration purposes. Image taken from [60].

The online calibration approach presents some advantages in that:

1. No calibration object is required to be imaged after the clinical intervention is performed.
2. Only need rotate the X-ray source to acquire two views.
3. Use angiographic image point correspondences to recover both the internal and external gantry parameters.

More recently, techniques have been developed for determining the imaging geometry from the patient image data alone [57-58, 78-81, 82-84]. The basic principles of some of these more recent approaches will be discussed below. In each of the techniques, which do not require a calibration object, the SID (source-to-object distance), the pixel sizes, a scale factor, and corresponding points in the two images must be known. The approach presented in [57, 58] can approximate the values of the source-to-intensifier distance as well if a sufficient number of corresponding points is available. The scale factor is necessary for measurement of absolute distances and can be determined either by measurement of the distance between focal spots or by measurement on a calibration object (e.g., catheter markers or ball) imaged along with the patient or separately at the same gantry (imaging geometry) settings. The bifurcation points, i.e., the point of intersection of the centerlines of parent and branch vessels, are usually employed as the

points in the images, because their correspondence can be established uniquely from the hierarchical relationships in the vascular tree. Each of the proposed methods requires a minimum number of points (2 [57, 58], 5 [82], 8 [80, 81]). The iterative techniques, although different from one another, use the 2D image data for calculating the rotation and translation that minimize the differences between the input image data and the projections of the calculated 3D points into the image plane in a non-linear manner.

2.4.3 Feature Correspondence

Correspondence algorithms generally consist of two parts: a similarity measure and a cost function. The similarity measure is used to compare two points or regions in different images. The cost function analyzes the values produced by the similarity measure after a point or region in one image has been compared to many points or regions in the other image. In general, correspondences are found by choosing the point with the optimum similarity value, but this can give unsatisfactory correspondences. For example, in a grayscale image, a point in the middle of a region of texture is likely to match well to a large area, but may have as its optimum values, a point near the boundary of the region, rather than near the middle as would be expected. As a result, neighboring points will often have correspondences that are not close to each other. Some authors attempt to overcome these problems by smoothing the calculated correspondences so that the resulting displacement vectors all point in approximately the same direction [85], or detect and discard unreliable correspondences [86]; others grow correspondences from a distinctive point [87, 88]. Methods that involve segmentation often have reliable and robust correspondence algorithms [89-91], but apart from the difficulties of reliable segmentation, these methods cannot directly calculate correspondences for all the points in the image, and must interpolate correspondences for points not included in the segmentation.

2.4.3.1 Stereo Disparity Maps & Correspondence

Stereo matching is one of the most active research areas in computer vision and it serves as an important step in many applications (e.g., view synthesis, image based rendering, *etc*). The goal of stereo matching is to determine the disparity map between an image pair taken from the same scene. Disparity describes the difference in location of the corresponding pixels and it is often considered as a synonym for inverse depth (Figure 2.7). Due to the ill-posed nature of the stereo matching problem, the recovery of accurate disparity still remains challenging, especially in textureless regions, disparity discontinuous boundaries and occluded areas.

In general, stereo algorithms can be categorized into two major classes. The first class is local (window-based) algorithms, where the disparity at a given pixel depends only on intensity values within a finite neighboring window. Local methods can easily capture accurate disparity in highly textured regions, however they often tend to produce noisy disparities in textureless regions, blur the disparity discontinuous boundaries and fail at occluded areas. The second class is global algorithms, which make explicit smoothness assumptions of the disparity map and solve it through various minimization techniques. Recently, global methods such as graph cuts have attracted much attention due to their excellent experimental results.

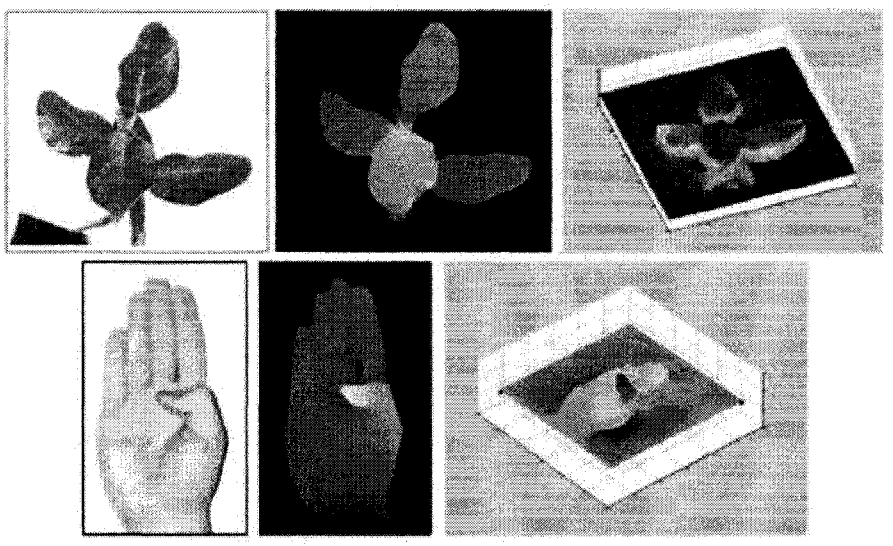


Fig 2.7. Sample depth maps and corresponding 3D plot from rigid images. Images taken from [92].

2.4.3.2 Epipolar Geometry Correspondence

A fundamental geometric correspondence method in the computer vision field relies on the epipolar geometry concepts. The epipolar geometry between two views is essentially the geometry of the intersection of the image planes. The baseline is the line joining the camera centers. The search for corresponding points between both images usually crafts this geometry. Suppose a 3D point X is imaged in two views, at x in the first, and x' in the second. This begs the question- what is the relation between the corresponding image points x and x'? To answer this, from (Figure 2.8 -left) we define the epipoles (e, e') as the points of intersection of the line joining the camera centers (the baseline) with the image plane. Equivalently, the epipole is the image in one view of the camera centre of the other view. An epipolar line is the intersection of an epipolar plane with the image plane. All epipolar lines (xe and $x'e'$) intersect at the epipole. An epipolar plane intersects the left and right image planes at epipolar lines that define the corresponding points. The fundamental matrix, F, is the algebraic representation of epipolar geometry. Given a pair of images, it is noted that for each point x in one image, there exists a

corresponding epipolar line ($x'e'$) in the other image. Any point x' in the second image matching the point x must lie on the epipolar line ($x'e'$).

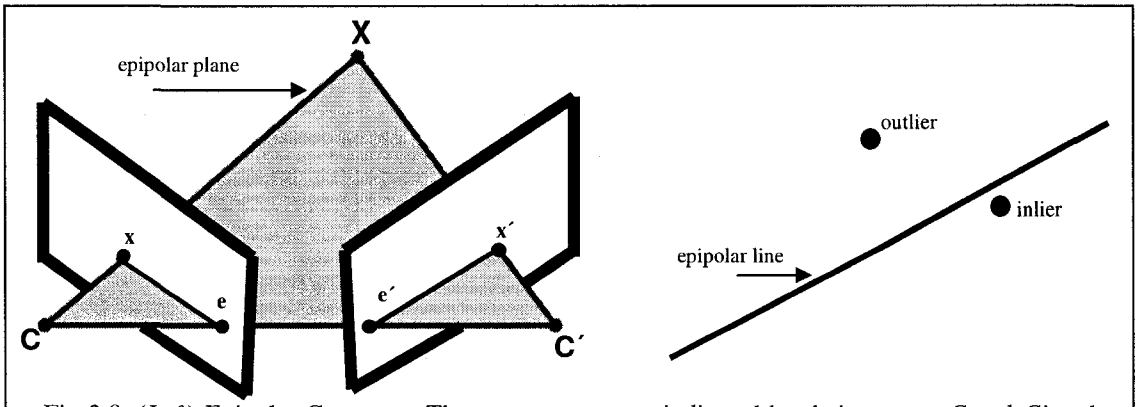


Fig 2.8. (*Left*) Epipolar Geometry. The two cameras are indicated by their centers C and C' and image planes. The camera centers, 3D point X, and its projected images x and x' lie in a common epipolar plane. (*Right*) Corresponding point matches between the two views can be classified as either outliers (mismatches) or inliers (correct matches), depending on their distance from the epipolar line.

The epipolar line is the projection in the second image of the ray from the point x through the camera centre C of the first camera. Thus, there is a map from a point in one image to its corresponding epipolar line in the other image. It will turn out that this mapping is a projective mapping from points to lines, which is represented by the fundamental matrix. In order to calculate the fundamental matrix, a normalized 8-point algorithm is used as described by Hartley [93]. If we define point x in the left image as $[u\ v\ 1]$ and point x' in the right image as $[u'\ v'\ 1]$ we then obtain the fundamental matrix using:

$$[u' \quad v' \quad 1]F\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0 \tag{2.9}$$

The fundamental matrix encompasses the geometrical information of the two camera setup. The effect of noise generates two potential types of errors during correspondence: an incorrect location of a pair (inlier error) match, and incorrect pairing (outlier error).

(Figure 2.8-right) shows these two types of errors. Inlier errors are assumed to display a Gaussian distribution. This means that most errors will be small and within one or two pixels.

Typically, if more than eight corresponding points are available, a least mean square minimization is often used. Hartley showed that to make the resulting fundamental matrix satisfy the rank two requirements, the singularity enforcement is performed [93]. Its simplicity of implementation is advantageous. Also, it is noted that after outlier (false matches) rejection, the eight point algorithm performs comparably with other nonlinear optimization techniques. One of these techniques relies on cross correlation, however false matches are unavoidable in this case. A more intricate nonlinear iterative optimization method was therefore proposed in [94]. To guide the optimization process, these nonlinear techniques use objective functions, such as the distance between points and their corresponding epipolar lines. Despite the increased robustness of these methods, nonlinear optimization methods require somewhat careful initialization for early convergence of the cost function. Torr and Murray [95] proposed RANSAC (random sample consensus), which randomly samples a minimum subset with seven pairs of matching points for parameter estimation. The candidate subset that maximizes the number of points and minimizes the residual represents the solution. Nevertheless, it is computationally infeasible to consider all possible subsets, which tend to be exponential in number. Therefore, additional statistical measures are needed to derive the minimum number of sample subsets. RANSAC is considered to be one of the most robust methods for this.

In the clinical community, common practice for biplane point correspondence relies heavily on the epipolar constraint. Schechter et al. relates the matching problem to finding the optimal path through a correspondence matrix representing all possible correspondences between the two images used [96]. In [97-99], the authors again rely on the epipolar planes and epipolar lines in order to trace corresponding coronary vessel segments, but with a slight difference in that an additional third image is used (i.e. three

view reconstruction), since among the several matching candidates given by the epipolar constraint, only one will be coherent with the additional view in most cases [99].

2.4.4 3D Reconstruction

Spatial reconstructions from biplane angiograms evolved as important tools for morphological analyses of vessel trees in both cardiology and neurology domains [94-99]. From the known imaging geometry and based on the epipolar constraint, any point visible in both projections can be spatially reconstructed by retracing the projection rays back to the point of their intersection. This can be achieved after successful calibration of the X-ray system. Since the reconstructed rays often miss this point due to slight reconstruction or calibration errors, usually their closest location is estimated as approximation. This is known as a stereo triangulation procedure (Figure 2.9).

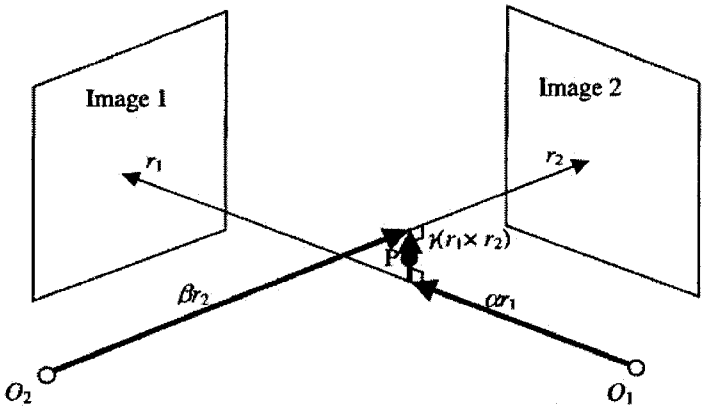


Fig 2.9. Stereo triangulation setup for the 3D reconstruction of corresponding points located in biplane images. Since both backprojected rays from camera O_1 and camera O_2 do not intersect, a solution is found by estimating a 3D point, P, which is half way between the two rays.

More than two views may be used as well for the reconstruction process. For example, Venaille proposed acquiring a third projection to reduce the number of consistent

matches between arterial segments [100]. This means that the first two images are used to obtain a 3D structure. Then, the 3D structure is reprojected back onto the third image. Only those image points that are compatible with the reprojection are used as correspondences between the three views.

2.4.5 Biplane Reconstruction Limitations

All biplane algorithms developed in the clinical context for the 3D reconstruction of coronary arteries published 3D reconstruction root mean square errors (RMS) which are less than 2 mm using a single pair of perpendicular views. To self calibrate the biplane systems initial methods used a calibration object imaged at the end of the angiographic acquisition. Later, methods exploited the corresponding landmarks between a single pair of biplane images to estimate the X-ray gantry parameters in order to reduce reconstruction errors. Cheriet et al. [101] demonstrated that self-calibration and 3D reconstruction of the coronary arteries becomes possible using a spatio-temporal technique that uses as landmarks the coronary bifurcation points. For a minimal number of point matches, approximately 30 corresponding bifurcation points across 5 consecutive images, the root mean square error remains stable. However, the study also demonstrated that the 3D RMS error will rise as the total number of images increases since we are systematically introducing spatial errors for reconstruction purposes. Lastly, we cannot rely on bifurcation point correspondence as these may be limited in number in the angiographic images. Therefore, it is essential to extract other coronary landmarks to perform calibration and reconstruction.

In computer vision literature, the RANSAC method is primarily used for point correspondence of rigid moving objects, more specifically for those that have line-like structures or lie on dominant planes in the computer vision context. The method randomly selects which point matches to use when performing 3D reconstruction of a scene. To use the RANSAC method to filter out all possible point matches when considering the coronary arteries would be a mistake since these undergo non rigid movements during the heart contractions. Therefore, another method should be

considered to select those point matches that undergo non rigid motion and that would yield acceptable reconstruction errors.

2.5 Single View (Monoplane) Reconstruction

Monocular 3D reconstruction is inherently an ambiguous problem, but using prior knowledge about a domain, it is often possible to recover distances and spatial configurations from a single image or a set of monoplane images. We survey the pertinent algorithms in the fields of computer vision and in the clinical context.

2.5.1 3D Vision Community

Most reconstructions are only determined up to an arbitrary projective transformation. This might be sufficient for some robotics applications, but certainly not for visualization. Hence, a method is needed to upgrade the reconstruction to a metric one (i.e. determined up to an arbitrary Euclidean transformation and a scale factor). In general three types of constraints can be applied to achieve this: scene constraints, camera motion constraints and constraints on the camera intrinsic parameters. All of these have been tried separately or in conjunction. In the case of a hand-held camera and an unknown scene, only the last type of constraints can be used. Reducing the ambiguity on the reconstruction by imposing restrictions on the intrinsic camera parameters is termed *self-calibration* (in the area of computer vision) [102].

For long monoplane video sequences, the accumulation error in the projective reconstruction may ultimately cause the self calibration to fail. When obtaining a 3D reconstruction from long image sequences, a subset of frames should be selected that are most suitable for the estimation of the epipolar geometry between views. This is advantageous since it is impractical to use all image frames contained in a sequence. The frames belonging to this subset are called keyframes. In order to accurately estimate the camera motion between single view image frames, key-frames should be selected with a sufficient baseline (disparity or distance) between them. If two consecutive key-frames

are too close to each other some of the degrees of freedom of the epipolar geometry are hard to estimate which leads to problems in motion estimation and in identifying suitable feature points [102].

The second issue that needs to be addressed when processing long image sequences is the mitigation of projective drift. Temporal structure and motion recovery is incrementally constructed by orienting additional views using the 3D positions of common features. Inaccuracies in the 3D positions of feature points due to noise and approximation errors accumulate over the course of the sequence and have a detrimental effect on the overall reconstruction. What this means is that as the number of images are added towards the reconstruction process, we inherently add additional possible errors related to the spatial configuration of the image coordinates and 3D structure. Projective drift is most problematic for self-calibration methods over long sequences, but has little effect over short sequences [102].

Pollefey and al. present an approach that is able to reconstruct 3D models from extended video sequences captured with an uncalibrated hand-held camera [102]. The focus is on two specific issues: (i) key-frame selection, and (ii) projective drift. Given a long video sequence it is often not practical to work with all video frames. In addition, to allow for effective outlier rejection and motion estimation it is necessary to have a sufficient baseline between frames. For this purpose, the authors propose a key-frame selection procedure based on a robust model selection criterion. The approach guarantees that the camera motion can be estimated reliably by analyzing the feature correspondences between three consecutive views. After self calibration, triplets of key-frames are aligned using absolute orientation and hierarchically merged into a complete metric reconstruction. A detailed 3D surface model using stereo matching is constructed in the last step.

The Geometric Robust Information Criterion (GRIC) model selection approach is useful in key-frame selection. The GRIC computes a score based on the feature correspondences between images and a model. The key-frame criterion proposed by the authors selects three views where the GRIC score of the epipolar model (F) is lower than

the score of the homography (H) model. The frames are processed sequentially, and the first view in the sequence is selected as the first key-frame. Subsequent views are selected as the keyframe if the fundamental matrix between it and the last keyframe view yields a GRIC(F) score that is smaller than the GRIC(H) score of the homography between the two views. The features used in computing the GRIC are those from the current view that are found in the previous two key-frames. By using features found in the current view and the previous two key-frame views, the authors ensure that the three-view geometry can be estimated reliably for each selected triplet of views [102]. Figure 2.10 shows the reconstruction process developed by Pollefey.

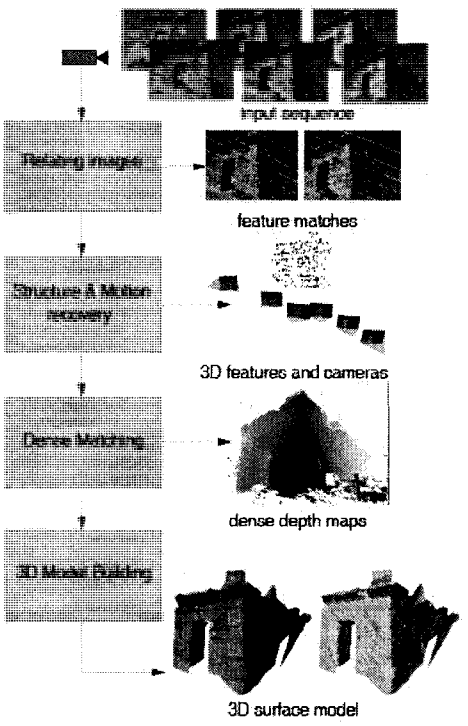


Fig 2.10. Overview of the reconstruction process. Image taken from [102].

1- An input sequence of images is selected and the GRIC scores permit to select the first three images for reconstruction.
2- The next step attempts to match all image pixels of an image with pixels in neighboring images, so that these points can also be reconstructed.
3- This task is greatly facilitated by the knowledge of all the camera parameters that were obtained in the previous stage. Since a pixel in the image corresponds to a ray in space and the projection of this ray in other images can be predicted from the recovered pose and calibration, the search of a corresponding pixel in other images can be restricted to a single line.
4- A depth estimate (i.e. the distance from the camera to the object surface) for almost every pixel of an image can be obtained. A complete dense 3D surface model is obtained by fusing the results of all the depth images together.
5- The images used for the reconstruction can also be used for texture mapping so that a final photo-realistic result is achieved.

In [103], the authors present a method for object-based rigid 3D motion estimation from monoscopic image sequences. Initial 2D vectors are available at the projections of the 3D nodes on all the image planes for multiview camera geometry. The rigid 3D motion

of each articulated object in the scene is estimated using a neural network based on the available 2D motion information on the image planes of the multiview camera geometry. The weights of the neural network are updated using nonlinear error minimization techniques. The technique develops a novel approach for flexible 3D motion estimation of each node of the object model. The performance of the rigid and non-rigid 3D motion estimation techniques is evaluated experimentally on both synthetic and real 3D object motion. For the monocular case, only the rigid motion is recovered with the supposition that the 3D coordinates are known at each time instant. It has as objective to recuperate the camera parameters accurately and then to perform 3D reconstruction using the optimal parameters. Monoplane reconstruction results showed that for the first two time instants, the mean displacement error for the 3D rigid motion was 8.1%.

In [104], authors showed how a mild assumption about an indoor scenes' geometry and efficient floor boundary detection can be used to recover fairly accurately 3D information from a single image. The approach encodes prior knowledge about the environment such that the floor is flat and that walls are vertical, and is based on recognizing the floor and the wall regions in the image.

Methods for single image reconstruction commonly use cues such as shading, silhouette shapes, texture, and vanishing points [105-109]. These methods restrict the allowable reconstructions by placing constraints on the properties of reconstructed objects (e.g., reflectance properties, viewing conditions, and symmetry). A few approaches explicitly use examples to guide the reconstruction process. One approach [110, 111] reconstructs outdoor scenes assuming they can be labeled as "ground," "sky," and "vertical" billboards. A second notable approach makes the assumption that all 3D objects in the class being modeled lie in a linear space spanned using a few basis objects (e.g., [112-115]). This approach is applicable to faces, but it is less clear how to extend it to more variable classes because it requires dense correspondences between surface points across examples. In [116], the researchers use a database of objects from single class (e.g. hands, human figures) containing example patches of feasible mappings from the appearance to the depth of each object. Given an image of a novel object, they

combine the known depths of patches from similar objects to produce a plausible depth estimate. This is achieved by optimizing a global target function representing the likelihood of the candidate depth. They demonstrate how the variability of 3D shapes and their poses can be handled by updating the example database instantly.

2.5.2 Clinical Community

Only two single view algorithms are proposed in the cardiology clinical setting worldwide that do not rely on registration. The clinical domains targeted are the intravascular ultrasound interventions (IVUS) and the catheter ablation procedures. The mathematical formulations for both will be presented below.

2.5.2.1 IVUS Transducer Single Plane Reconstruction

During an IVUS intervention, a catheter with an ultrasound transducer is introduced in the body through a blood vessel and then pulled back to image a sequence of vessel cross-sections. Unfortunately, there is no 3D information about the position and orientation of these cross-section planes. To position the IVUS images in space, some researchers have proposed complex stereoscopic procedures relying on biplane angiography to get two X-ray image sequences of the IVUS transducer trajectory along the catheter. Sherknies et al. [117] and Jourdain et al. [118] have elaborated a much simpler algorithm to recover the transducer 3D trajectory with only a single view X-ray image sequence. The known pullback distance of the transducer (0.1cm/sec) during the IVUS intervention is used as an a priori to perform this task. The pullback value gives the true 3D displacement of the transducer in the world coordinate system. Considering that biplane systems are difficult to operate and rather expensive and uncommon in hospitals, this simple pose estimation algorithm could lead to an affordable and useful tool to better assess the 3D shape of vessels investigated with IVUS.

Sherknies et al. first introduced the concept of a monoplane approach when projecting the 2D image position of the transducer back into 3D spatial coordinates. Two

types of projection geometry were analyzed: a weak-perspective and full-perspective model (Figure 2.11).

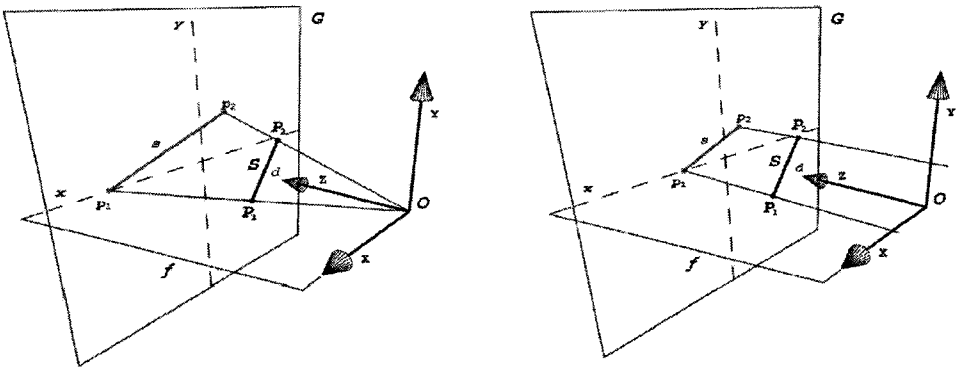


Fig 2.11. (*Left*) Full perspective projection geometry. Segment *S*, is projected onto the plane *G*. All projectors originate from the origin, *O*. (*Right*) Weak perspective projection model. All projectors are parallel towards the plane. Image taken from [117].

By defining a 3D spatial point as $P_i = [X_i, Y_i, Z_i]$, and similarly a 2D image point as $p_i = [x_i, y_i]$, the relationship between these points in a pin hole camera model is given by:

$$x_i = f\frac{X_i}{Z_i} \tag{2.10}$$

$$y_i = f\frac{Y_i}{Z_i} \tag{2.11}$$

where f is the focal distance. The authors assume that the projection plane is perpendicular to the Z axis and centered on it. By using the equation of the Euclidean distance between two points $d = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2}$ and rewriting equation (2.10) we have:

$$Z_2 = f\frac{X_2}{x_2} \tag{2.12}$$

We can now substitute equation (2.12) into equation (2.11) to arrive at:

$$Y_2 = y_2\frac{X_2}{x_2} \tag{2.13}$$

Substituting the above two equations into the Euclidean distance formula and solving for X_2, we obtain:

$$X_2 = \frac{x_2 a \pm \sqrt{x_2^2(a+bc)}}{b} \tag{2.14}$$

where

$$\begin{aligned} a &= x_2X_1 + y_2Y_1 + f\,Z_1 \\ b &= f^2 + x_2^2 + y_2^2 \\ c &= d^2 - X_1^2 - Y_1^2 - Z_1^2 \end{aligned}$$

Knowing the location of two projected image points, p_1 and p_2, the location of a 3-D point P_1 and the distance d between the two 3-D points, 3-D coordinates of P_2 can be calculated. The abscissa is found using equation (2.14), and by substituting this equation in equations (2.12) and (2.13) the depth coordinates are found. However, there are two possible solutions for the depth as a square root term is involved. The authors use an *a priori* knowledge of the experimental setup to resolve this issue. In weak-perspective projection geometry, the catheter path can be reconstructed without prior transducer depth information. The projection of each 3-D point is not affected by their respective depth coordinates but rather by the average scene depth. Hence, a 2D image point is related to its 3D spatial point by the following:

$$x_i = f'X_i \tag{2.15}$$

$$y_i = f'Y_i \tag{2.16}$$

where $f' = (\frac{f}{\overline{Z}})$ and $\overline{Z} = (\frac{1}{n})\sum_{i=1}^{n} Z_i$, and n is the number of points. Using the two above equations we can rewrite the Euclidean distance equation as:

$$Z_2 = Z_1 \pm \sqrt{d^2 - \frac{[(x_2 - x_1)^2 + (y_2 - y_1)^2}{f'^2}} \tag{2.17}$$

which gives the depth coordinate of point P_2. Results showed that for the weak perspective projection model, an average error of 5.4 mm was obtained for the depth estimation with a maximum error of 9 mm, whereas for the full perspective model the average depth estimation was 2.3 mm with a maximum depth of 5.7 mm.

Jourdain et al. continued the work of Sherknies in order to prune the two possible solutions for the depth estimation [118]. To prune the trajectories as much as possible during the depth estimation process, the authors have elaborated a cost function that gives a smaller weight to the smoothest solutions considering the number of undulations in the trajectory (Figure 2.12). The results of the method showed that the maximum RMS error in all the reconstructions for both phantoms is 3 mm, and the minimum RMS error was 0.1 mm depending of the amount of noise in the 2D transducer positions.

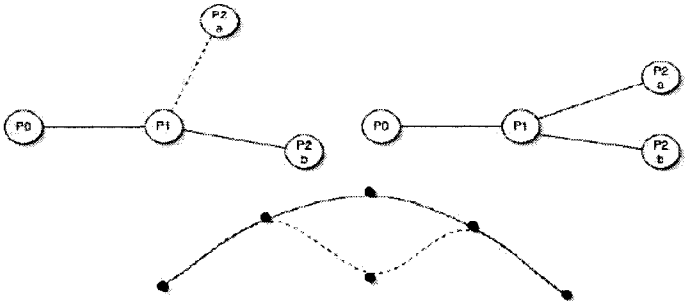


Fig 2.12. (*Top left*) Trajectory pruning based on curvature. The dashed trajectory can be rejected. (*Top right*) Here both solutions are acceptable. (*Bottom*) The cost function used will most likely favor the solid line configuration and reject the dashed line configuration for the transducer path that is reconstructed. Image taken from [118].

2.5.2.2 Catheter Ablation Single Plane Depth Estimation

Severe disorders of the heart rhythm that can cause syncope or sudden cardiac death (SCD), can be treated by radio-frequency (RF) catheter ablation. Fast cardiac rhythm, called tachycardia, is a common form of SCD. The precise localization of the arrhythmogenic site and the positioning of the RF catheter over that site are problematic: they can impair the efficiency of the procedure and are time consuming. Radio-frequency catheter ablation (RF ablation) consists of inserting a catheter inside the heart, near the area from which originates the abnormal cardiac electrical activity, and delivering RF currents (500 kHz, 10-25 W) through the catheter tip so as to ablate the arrhythmogenic area. This procedure is guided by X-ray images of the catheters obtained with a fluoroscopic system. Several commercial systems that provide a 3D color display of the cardiac electrical activation sequence have been applied in recent years to facilitate the mapping of complex arrhythmias however they are expensive. Fallavollita and al. [119] have been the first to elaborate an algorithm to estimate the depth of the ablation catheter using a single X-ray image.

The authors deduced that the inverse of the projection width of the ablation catheter tip is proportional to the depth of the catheter. The width of the catheter tip was determined by projecting the grayscale image along a line perpendicular to the electrode tip. This projection results in a bell-shaped curve (Figure 2.13) from which the tip electrode width can be measured.

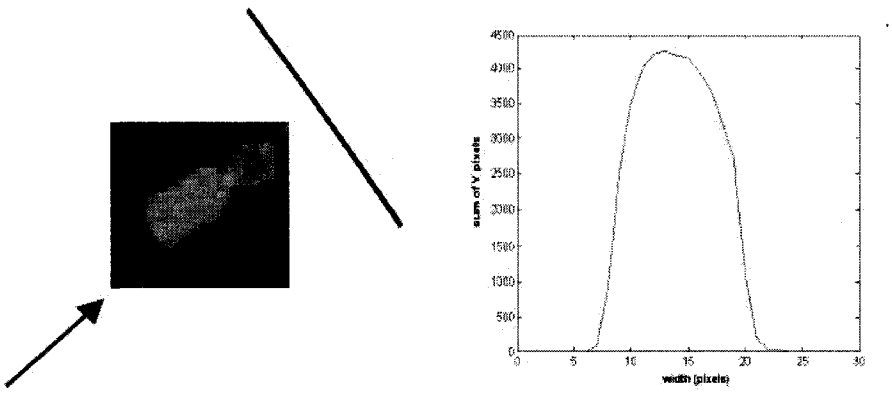


Fig 2.13. (*Left*) Cropped image of the ablation catheter obtained after image processing. (*Right*) The image was projected along a line perpendicular to the electrode tip yielding a bell shaped curve from which the width was approximated at the half way point. Image taken from [120].

Results showed that by using an orthographic projection model (image pixel coordinates backprojected perpendicularly onto the world frame) the error in depth estimation was about 10 mm. Further, the correlation coefficient was statistically significant (r = 0.66) when relating the depth to the inverse of the catheter width.

2.5.3 Monoplane 3D Reconstruction Limitations

The survey of single plane reconstruction algorithms in the computer vision community demonstrate again that objects of interest must be of rigid nature, or undergo rigid motion. The implementation of 2D stereo maps fails if the objects undergo large motion or are occluded by other objects in the scene. Other techniques make assumptions of the

scene geometry in order to arrive at a correct reconstruction. Shading, silhouette shapes, texture are a few of these cues used to infer depth estimation.

When considering the IVUS monoplane reconstruction algorithm, the points which are projected back into 3D space are the transducer positions. It is noted that the transducer is tracked at each image frame. Sherknies et al. analyzed a total of 821 and 818 image frames respectively when experimenting on synthetic images. Jourdain et al. made no mention of the amount of frames used to analyze the procedure. The authors used a *priori* information which was the pullback speed of the transducer (0.1cm/sec). This is extremely valuable when optimizing and solving for the estimated depth positions as the true 3D distances travelled by the transducer between image frames are known. When considering the arrhythmia ablation case, only a single image is considered for the catheter tip depth estimation. Fallavollita et al. use as information the width of the catheter tip, which is proportional to the unknown depth. However this value may be inaccurate due to the amount of noise present in an X-ray image and will lead to severe estimates with respect to the true depth value. Further, the actual 2D spatial image position of the catheter tip may be such that it is not fully flat but inclined to the viewer if not positioned properly on the arrhythmogenic site. This inclination will undervalue the width calculated and can produce incorrect depth estimates.

The authors above use only the focal length as the only C-arm parameter which is still considered an approximation when recovering 3D structure. Therefore a more complete calibration technique must be considered that relates the projection of the 3D structure onto the 2D images.

Chapter 3- Research Hypothesis and Objectives

Our research study will investigate the feasibility of incorporating image processing and reconstruction techniques using only a single monoplane angiographic image and a 3D *a priori* model for the sole purpose of clinical interventional assistance. We deem this important as these angiographic images are acquired using the primary imaging modality when treating cardiovascular disease such as atherosclerosis. The techniques developed can be immediately adapted to a secondary application for treating electrophysiological defects inherent with arrhythmias.

Interventional assistance signifies providing the cardiologist with some sort of additional information and help that will enable them to quicken diagnostic, quicken surgical methodology during the intervention, diminish overall acquisition and procedure time which in turn reduces X-ray dosage for both the patient and clinician (Figure 3.1). Before outlining our research hypothesis and specific objectives, we will first present the necessary tools required to meet our major research contributions.

3.1 Angiographic Image Segmentation and Temporal Tracking

This section describes the ideas required to arrive at a visual geometrical reconstruction of the arteries. The goal here is to target a method that requires no change or continuous selection in algorithmic parameters for the extraction of coronary arteries, like the Frangi filter. Ideally, we would want the clinician to provide a specific angiographic image, one that describes the diastolic cardiac cycle phase, and the algorithm would automatically enhance the principal coronary arteries with no fine tuning of parameters. Hence, we propose to develop a novel automatic filter that has all its parameters fixed and would run smoothly on any given angiographic image as seen on the fluoroscopy screens during intervention. To minimize user interaction, a two-click technique to select and extract the centerline of a coronary artery that is being investigated by the clinician will be proposed. This is important as it assists the interventionist by providing a smaller

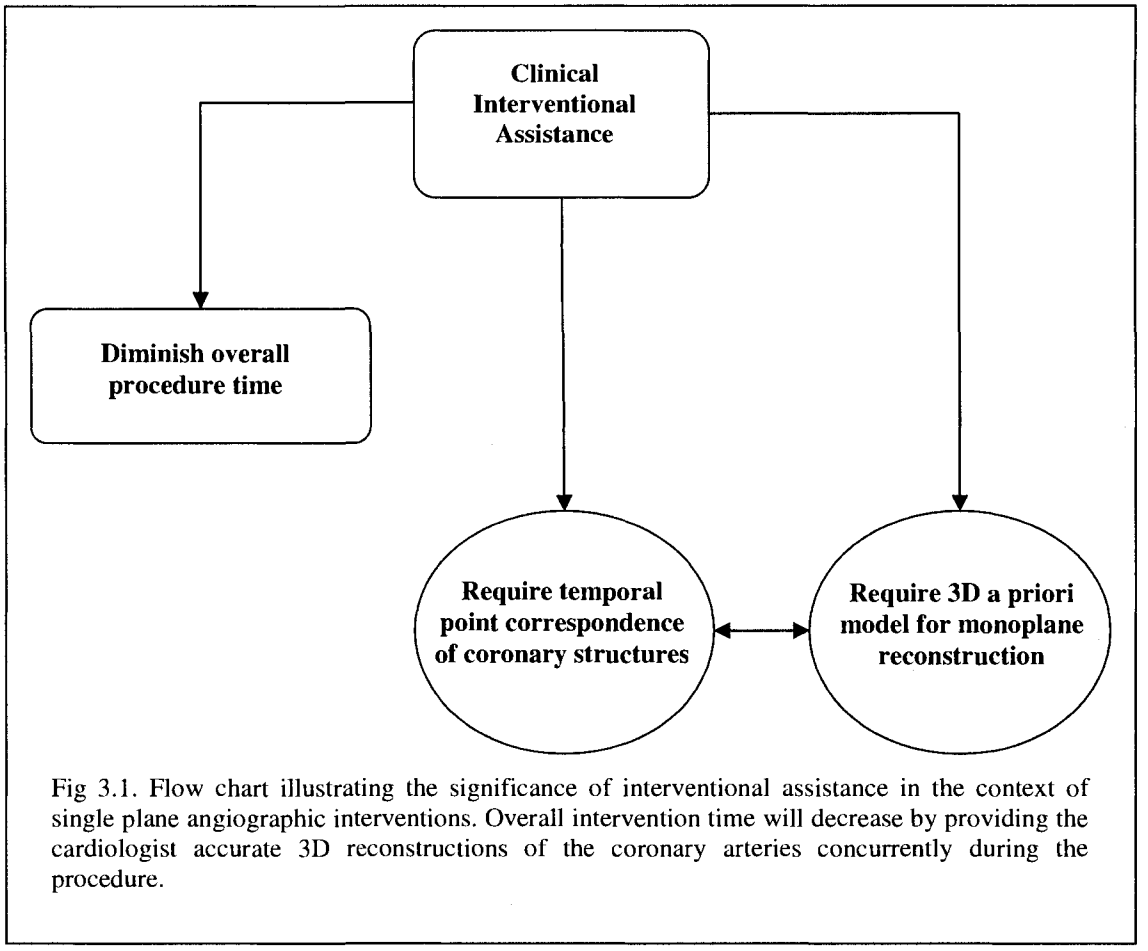


Fig 3.1. Flow chart illustrating the significance of interventional assistance in the context of single plane angiographic interventions. Overall intervention time will decrease by providing the cardiologist accurate 3D reconstructions of the coronary arteries concurrently during the procedure.

region of interest to analyze instead of surveying the entire angiographic image when trying to detect where a possible stenosis lies.

3.2 Optimal Point Correspondence for Non Simultaneous Images

Traditionally the potential corresponding points are chosen if they lie closer to their epipolar lines. To improve the overall selection process and refinement of point matches between two biplane images (Figure 3.2), we will introduce a geometric curvature criterion that takes into consideration the 2D and 3D curvature of the artery centerlines.

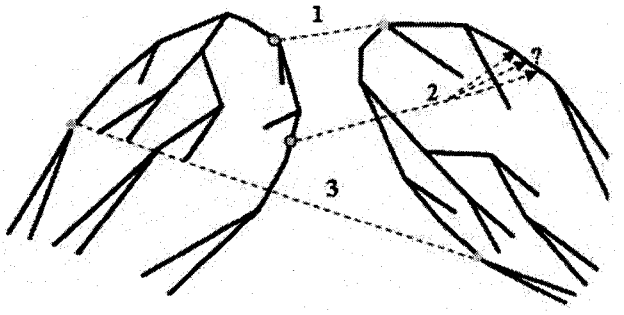


Fig 3.2. Biplane correspondence of coronary arteries becomes complicated when matching points, other than the bifurcations, between two views (line 2). When considering coronary bifurcation points as possible matches, the correspondence problem becomes elementary (lines 1 & 3). Image taken and adapted from [101].

Furthermore, the novel geometric constraint has several advantages. First, it allows us to target coronary image points that undergo only rigid motion. This particular constraint suggests that a coronary point that has undergone only rigid movement (non-contractility) will have the same Frenet-Serret curvature in both the 2D and 3D case. Moreover, since most relevant computer vision algorithms recover structure from motion by targeting only rigid structures, it would be appropriate to use the proposed curvature constraint in the cardiac context. Secondly, the curvature constraint will have the advantage of refining all possible point matches when considering a single view acquisition. This leads us to define our first research hypothesis.

Research Hypothesis 3.1: The integration of a geometrical curvature constraint improves the precision of the point correspondence and improves reconstruction accuracy.

Our proposition is inspired by the work in [101], however different in that every point on the principal coronary centerline will be attempted to be matched instead of using only the bifurcations points. This is important during a clinical intervention as a particular viewing angle of the X-ray fluoroscope may provide angiographic images that lack sufficiently in bifurcations.

3.3 Monoplane 3D Reconstruction

To assist the cardiologist in interventional procedures, we cannot make use of a two view setup if a single plane reconstruction algorithm can be developed. The point matching curvature criteria described above is primordial for monoplane reconstruction since optimal correspondences must be temporally tracked in a single angiographic sequence. By relating the 3D and 2D Frenet curvature information for this point, one can determine if it is a potential match before performing 3D reconstruction.

We propose to use *a priori* 3D information from pre-operative images in order to obtain a 3D representation of the coronary arteries at a first time instant. This 3D information is readily available to us before a specific angiographic operation. It can be obtained either by segmenting CT or MRI images before an intervention or it can be obtained by first acquiring two C-arm non-simultaneous images to retrieve the 3D information. We believe this initial information will allow the estimation of the 3D structure in subsequent X-ray images. This leads us to define our second research hypothesis.

Research Hypothesis 3.2: It is possible to recuperate the 3D geometry across a single plane angiographic sequence, for clinical assistance, by using an *a priori* 3D model of the structure of interest and tracking the artery centerlines across the sequence.

The strategy proposed here will primarily target the self-calibration of the gantry settings of the C-arm fluoroscope. Once the gantry parameters are self-calibrated to new optimal values, we will eventually target the 3D reconstruction of coronary artery segments that undergo strict non-contraction (rigid motion), such as the bifurcation points of parent coronary branches. Since estimating the 3D structure at later time instants involves optimizing a certain number of unknown variables and equations, we will consider analyzing a minimum number of X-ray images in order to estimate their 3D geometries. The method can be extended to reconstructing medical instruments such as ablation catheters for radio-frequency image guidance procedures as well.

The specific objectives addressed to meet the above hypothesis are as follows:

Research Hypothesis 3.1

3.1.1 Improve self calibration algorithms by introducing a novel curvature constraint that will help select valid 2D corresponding points between the images and make the reconstruction more precise.

3.1.2 Validate constraint on biplane images first to prove the necessity of refining the point matches in a single view sequence.

Research Hypothesis 3.2

3.2.1 Develop a novel automatic 2D segmentation algorithm to enhance healthy or diseased (stenosis) coronary arteries in a first image representing the diastolic phase.

3.2.2 Provide the interventionist with the ability to select a specific coronary artery and investigate its motion by extracting its 2D centerline and automatically temporal tracking it across all the angiographic images in a cardiac cycle.

3.2.3 Perform monoplane 3D reconstruction by implementing a novel self calibration algorithm using a single angiographic view and an *a priori* 3D model.

We conclude by re-emphasizing that our overall objective is to provide a clinical tool in order to assist the cardiologist throughout a specific intervention. The cardiologist uses an X-ray fluoroscope that is the primary modality in surgical rooms, and looks at fluoroscopy screens for guidance in angioplasty procedures. To help quicken the intervention, only specific optimal coronary points will be extracted and temporally tracked. We will exploit the rigid movement of these points to recover structure from motion. The optimal correspondences will provide a more accurate reconstruction for the artery being investigated and will lead to the development of a novel monoplane algorithm. The monoplane procedure makes use of a 3D model that reflects a specific time instant (i.e. diastolic cardiac phase). This *a priori* model serves as additional

information in order to estimate the 3D structure at later time instants in the X-ray fluoroscopy images.

Chapter 4- Methodology

The following chapter will highlight the necessary algorithms and tools that we propose in order to meet the specific objectives outlined in Chapter 3. The methods are to be applied using the X-ray angiographic images that are readily viewed on the fluoroscopy monitors. The sections described in this chapter include image segmentation, temporal tracking and ending with biplane and monoplane 3D reconstruction.

Shown in Figure 4.1, is a step by step depiction of the various algorithmic processes required in order to verify our two research hypotheses: optimal point correspondence and monoplane 3D reconstruction.

4.1 Automatic Image Extraction

In [120], an algorithm was developed to detect an optimal cardiac phase image frame. By calculating the standard deviation of the differences between all the pixels of two successive frames, we were able to determine that the frame showing the smallest standard deviation also had the least motion blur. This is to say that if we were to perform biplane 3D reconstruction, for example, the results would be more accurate if we analyze image frames where motion artifacts are minimal. This would occur when the heart motion is minimal and consequently when blurring is almost non-significant.

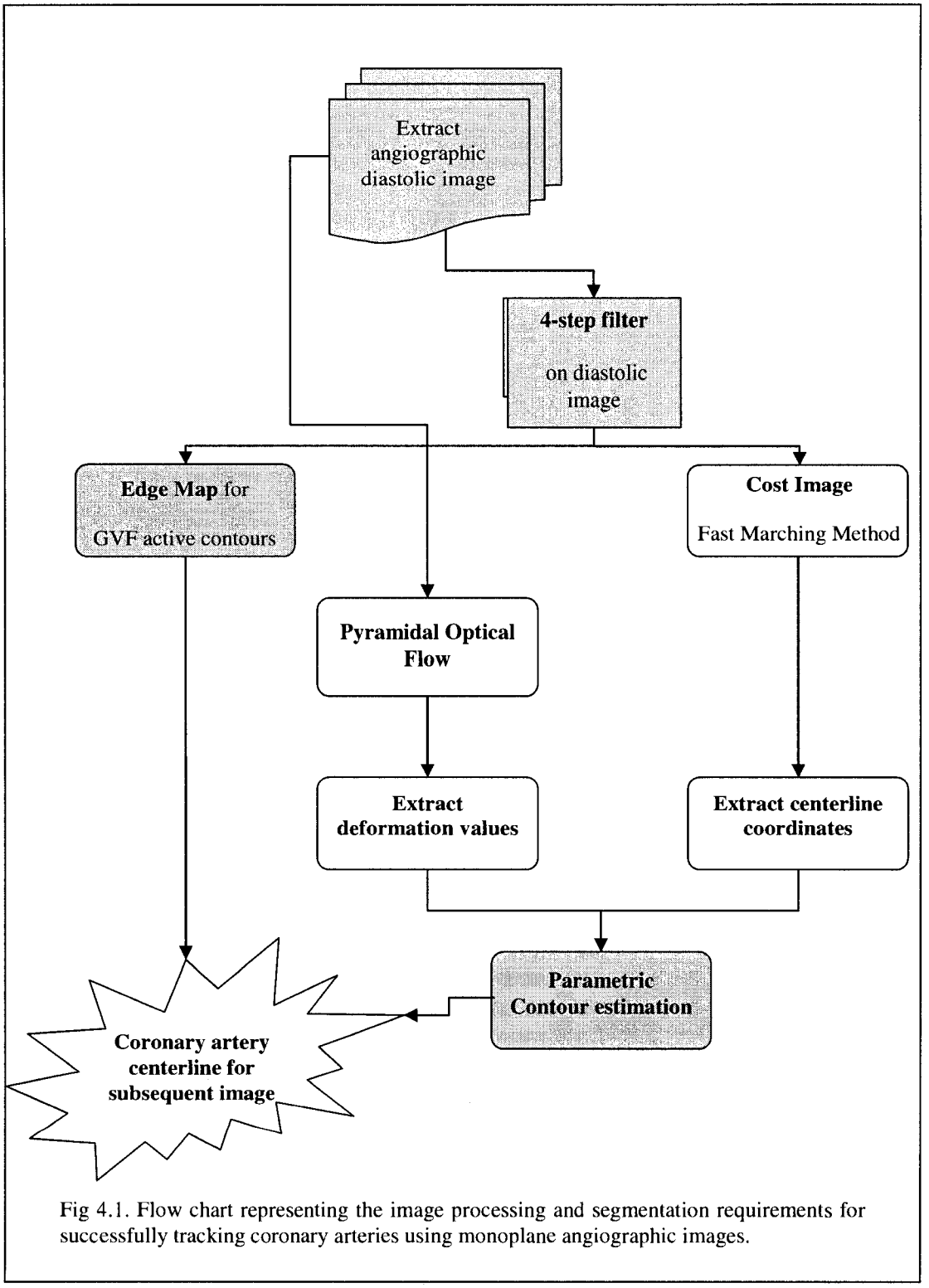


Fig 4.1. Flow chart representing the image processing and segmentation requirements for successfully tracking coronary arteries using monoplane angiographic images.

4.2 Image Preprocessing

Following is a brief description of the 4-step filter implemented in order to enhance the principal coronary arteries and suppress the background of the image.

4.2.1 Homomorphic Filtering

Frequency Domain involves the computation of the Fourier Transform of the image to be enhanced, multiply the result by a filter transfer function and take the inverse transform to produce the enhanced image [121].

4.2.1.1 High Pass Filtering

Edges and other abrupt changes in gray levels being associated with high frequency components, image sharpening can be achieved in the frequency domain by a high pass filtering process, which attenuates the low frequency components without disturbing high frequency information in the Fourier transform.

4.2.1.2 Low Pass Filtering

Edges and other sharp transitions (such as noise) in the gray levels of an image contribute significantly to the high frequency content of its Fourier Transform. Hence blurring is achieved in the frequency domain by attenuating a specified range of high frequency components in the transform of a given image. Given $F(u,v)$, the Fourier transform of the image to be smoothed, we need to find a transfer function $H(u,v)$ that yields $G(u,v)$ by attenuating the high frequency components of $F(u,v)$, in the equation,

$$G(u,v) = H(u,v) * F(u,v) \tag{4.1}$$

The illumination-reflectance model can be used as the basis for a frequency domain procedure that is used for improving the appearance of an image by simultaneous brightness range compression and contrast enhancement. An image $f(x,y)$ can be

expressed in terms of its illumination and reflectance components by means of the relation:

$$f(x,y) = \quad i(x,y)r(x,y) \tag{4.2}$$

The illumination component of an image is characterized by slow spatial variations, while the reflectance component tends to vary abruptly, particularly at the junctions of dissimilar objects. These characteristics lead to associating the low frequencies of the Fourier transform of the logarithm of an image with illumination and the high frequencies with reflectance. A good deal of control can be gained over the illumination and reflectance components with the homomorphic filter. This control requires specification of a filter function *H(u,v)* that affects the low and high frequency components of the Fourier transform in different ways. If the parameters γ_L and γ_H are chosen so that $\gamma_L < 1$ and $\gamma_H > 1$, the filter function shown tends to decrease the low frequencies and amplify the high frequencies [121]. The net result is simultaneous dynamic range compression and contrast enhancement. In conclusion, the homomorphic filter is given by:

$$H(u,v) = (\gamma_H - \gamma_L)\,(1 - e^{-c\cdot(D^2(u,v)/D_o^2)}) + \gamma_L \tag{4.3}$$

The coefficient c controls the sharpness of the slope at the transition between high and low frequencies, whereas D_o is a constant that controls the shape of the filter and *D(u,v)* is the distance in pixels from the origin of the filter [121].

4.2.2. Anisotropic Diffusion

Anisotropic systems are those that exhibit a preferential spreading direction while isotropic systems are those that have no preferences. The Perona-Malik anisotropic diffusion method [122] was implemented here in order to reduce and remove both noise and texture from the image, as well as, to preserve and enhance structures. The diffusion equation is given by

$$\frac{\partial I}{\partial t} = div(c(x,y,t)\nabla I) \tag{4.4}$$

Where I is the input image and $c(x, y, t)$, the diffusion coefficient, will control the degree of smoothing at each pixel point in the image. The diffusion coefficient is a monotonically decreasing function of the image gradient magnitude. It allows for locally adaptive diffusion strengths; edges are selectively smoothed or enhanced based on the evaluation of the diffusion function. Although any monotonically decreasing continuous function of the gradient would suffice as a diffusion function, two functions have been suggested:

$$\begin{aligned} c_1(x, y; t) &= \exp(-(\frac{|\nabla I|}{K})^2) \\ c_2(x, y; t) &= \frac{1}{1+(\frac{|\nabla I|}{K})^{1+\alpha}} \end{aligned} \tag{4.5}$$

The variable K is referred to as the *diffusion constant* or the *flow constant*. The variable α, controls the speed of diffusion and takes on a value greater than zero. Obviously, the behavior of the filter depends on K. The greatest flow is produced when the image gradient magnitude is close to the value of K. Therefore, by choosing K to correspond to gradient magnitudes produced by noise, the diffusion process can be used to reduce noise in images. Assuming an image contains no discontinuities, object edges can be enhanced by choosing a value of K slightly less than the gradient magnitude of the edges.

4.2.3. Complex Shock Filtering

Gilboa et al. have recently developed a filter coupling shock and linear diffusion in the discrete domain, showing that the process converges to a trivial constant steady state. To regularize the shock filter, the authors suggest adding a complex diffusion term and using the imaginary value as the controller for the direction of the flow instead of the second derivative. The complex shock filter is given by

$$I_t = -\frac{2}{\pi}\arctan\left(a\,\mathrm{Im}\left(\frac{I}{\theta}\right)\right)|\nabla I| + \lambda I_{\eta\eta} + \tilde{\lambda} I_{\xi\xi} \tag{4.6}$$

where a is a parameter that controls the sharpness of the slope, $\lambda = re^{i\theta}$ is a complex scalar, $\tilde{\lambda}$ is a real scalar, ξ is the direction perpendicular to the gradient and η is the direction of the gradient. The complex shock filter results in a robust and stable deblurring process that can still be effective in noisy environments, such as in fluoroscopy, due to the low signal to noise ratio of these images. The complex filter is an elegant way to avoid the need of convolving the signal in each iteration and still get smoothed estimations. The time dependency of the process is inherent, without the need to explicitly use the evolution time t. Moreover, the imaginary value receives feedback - it is smoothed by the diffusion and enhanced at sharp transitions by the shock, thus can serve better for controlling the process than a simple second derivative. A complete derivation is found in [123].

4.2.4. Morphological Operation

Morphological filtering was applied as a final image processing step in order to eliminate background elements around the primary coronary arteries. The structuring element consists of a pattern specified as the coordinates of a number of discrete points relative to a defined origin. Normally Cartesian coordinates are used and so a convenient way of representing the element is as a small image on a rectangular grid. We chose a disk structuring element that has a radius of a few pixels, since the contours of the coronary arteries can be modeled as a collection of disks varying slowly at a certain radius, and its centerline is the union of the centers of these disks. Figure 4.2 shows an example of such a structuring element. The origin is marked by a circle around that point. The structuring element will suppress the background (black) and enhance the arteries (grayscale). When a morphological operation is carried out, the origin of the structuring element is typically translated to each pixel position in the image in turn, and

then the points within the translated structuring element are compared with the underlying image pixel values [124].

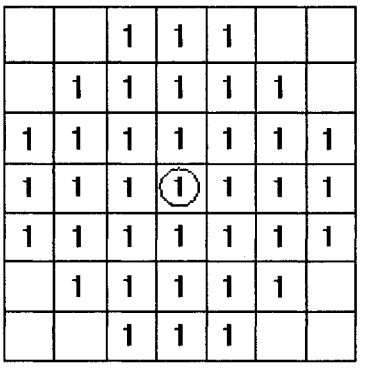


Fig 4.2. Shown is an example of a disk structuring element of radius 3 pixels. The origin pixel is circled and all other vacant pixel values are set to zero.

4.3 Centerline Extraction

In order to perform temporal tracking and 3D reconstruction of the obstructed coronary arteries, we need to extract its centerline in a first image representing the diastolic phase. A two-click centerline extraction approach will be introduced in the following section.

4.3.1 Fast Marching Method

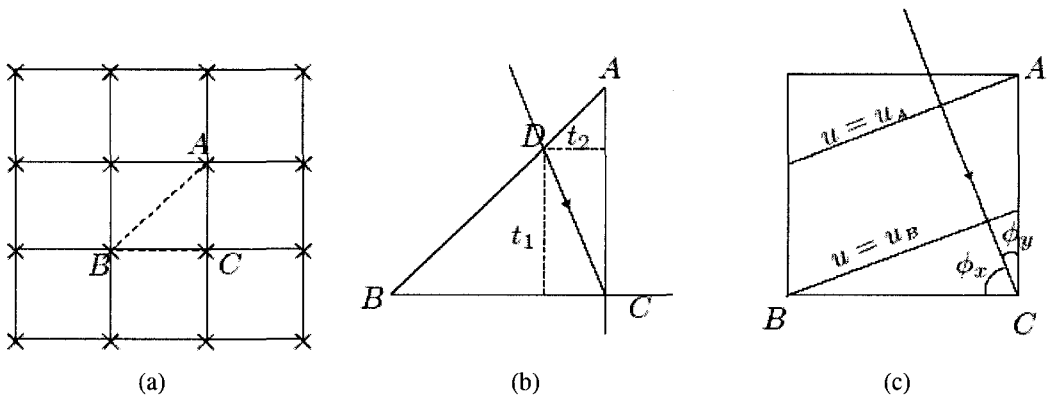
The proposed method for extracting the centerline of the artery is the Fast Marching Method. The isotropic form of the fast marching method algorithm is inspired by Qingfen Lin's thesis [30]. In Cartesian coordinates, the goal is to find the shortest path between two points that minimizes a cost function. If the cost τ is only a function of the location x in the image domain, then the cost function is called isotropic. Hence, if the cost function is isotropic, the arrival time map U satisfies the Eikonal equation

$$\|\nabla U\| = \tau \tag{4.7}$$

The discrete values of the above equation are of interest. On a 2D Cartesian grid with coordinates i and j, it was shown by previous authors that the solution to the following equation will give the correct solution u_{ij} for U at (i, j):

$$\max\left(u_{i,j}-u_{i-1,j},\, u_{i,j}-u_{i+1,j},0\right)^2+\max\left(u_{i,j}-u_{i,j-1},u_{i,j}-u_{i,j+1},0\right)^2 \quad =\tau_{i,j}^2 \qquad (4.8)$$

We recall that a single-pass algorithm to solve equation (4.7) was proposed independently by Sethian [27] who named it *Fast Marching*. The principal strategy of this algorithm is to introduce an order in the selection of the grid points of a 2D coordinate system, in a way similar to the well-known Dijkstra algorithm. This order is based on the fact that the arrival time u at any point depends only on neighbors that have smaller values. Thus, one-sided differences in equation (4.8) that look in the "upwind direction of the moving front" are sufficient when solving equation (4.7). In the 2D Cartesian space shown in Figure 4.3, the cost function is defined at the cross-positions in the center of each pixel. Assume that towards the point C, a planar wave is arriving in the northwest quadrant as in Figure 4.3c.



(a) (b) (c)

Fig 4.3. (*a*) 2D Cartesian grid. The arrival time at point C is to be computed from point A and B. (*b*) The arrival time at point C is approximated as the summation of the arrival time at point D and the time of traveling from D to C. (*c*) Alternative interpretation of (*b*) given by the direction cosines.

Since the farthest way in traveling in the isotropic case always occurs along the direction that is perpendicular to the wave front, the arrival time u_c at point C satisfies:

$$\begin{aligned} u_C - u_A &= \tau_C \cos\phi_y \\ u_C - u_B &= \tau_C \cos\phi_x \end{aligned} \tag{4.9}$$

Since ϕ_x and ϕ_y are complementary angles, we have

$$(u_C - u_A)^2 + (u_C - u_B)^2 = \tau_C^2 \tag{4.10}$$

which is one of the solutions to the quadratic formula of equation (4.8). The fast marching method is illustrated in Figure 4.4 and summarized in Outline 4.3.1. The method resembles Dijkstra's shortest path algorithm and is described as follows. Since $u(x_m)$ is the smallest among the not yet accepted points, and u-values only depend on points that have smaller u-values than itself, $u(x_m)$ will not be changed further and can be moved to the Accepted set. The updating step after accepting point x_m is given in Lin's thesis and omitted here for brevity [30]. In the updating scheme given inside the box, new values are only computed from already accepted points. This is an efficient implementation since unnecessary calculations are excluded. To determine the minimum-cost path after the end point is reached, a back-propagation from the end point to the starting point is carried out. This proposed method for centerline extraction limits user interactivity with the image by clicking only 2 points for initialization (start and end point) [30]. To initialize the algorithm we chose as a cost function the inverse of the image obtained by the 4-step filter.

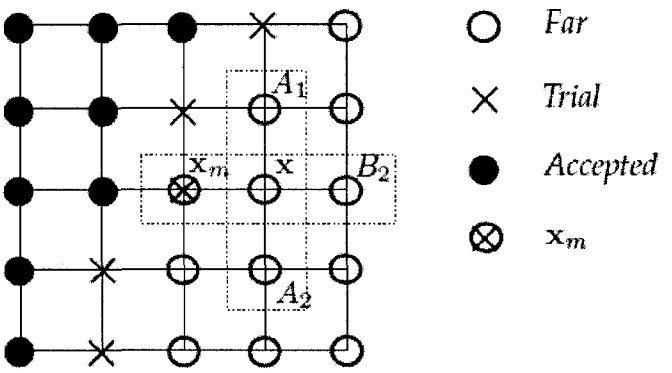


Fig 4.4. The Fast Marching Method. The point x_m has the minimum arrival time among all the Trial points. It will be moved to the Accepted set and the arrival times of its neighbors are to be updated.

Outline 4.3.1: Fast Marching Algorithm

Definition:
1. ***Accepted*** is the set of all grid points at which the u-value will not be changed.
2. ***Trial*** is the set of grid points to be examined. Their u-values have been computed.
3. ***Far*** is the set of all the other grid points. Their u-values have never been computed.
4. The cost function $\tau(x)$ is the cost of traveling at each grid point.

Initialization:
5. ***Trial*** set is the starting point $x0$, with $u(x0) = 0$;
6. ***Far*** set is all the other grid points with $u(x) = 1$;

Loop:
7. Let x_m be the ***Trial*** point with the smallest u-value;
8. Move it from the ***Trial*** set to the ***Accepted*** set;
9. Update neighbors of x_m

4.4 Single View Tracking Algorithm

Following the centerline extraction procedure, we are now ready to temporally estimate the positions of the artery centerlines in subsequent image frames. The tools required to perform such a task are described below.

4.4.1 Pyramidal Lucas-Kanade Optical Flow

The traditional Lucas-Kanade optic flow technique proposes to solve the brightness constant constraint equation by assuming constant flow over a fixed neighborhood region, w:

$$\begin{bmatrix} \sum wI_x^2 & \sum wI_xI_y \\ \sum wI_xI_y & \sum wI_y^2 \end{bmatrix} \begin{pmatrix} u_{flow} \\ v_{flow} \end{pmatrix} = -\begin{pmatrix} \sum wI_xI_t \\ \sum wI_yI_t \end{pmatrix} \tag{4.11}$$

Typically, the optical flow solution is solved by defining a small [5x5] window region, w and assigning a high weight to the center pixel. The gradients of the image I_x, I_y, as well as, the temporal gradient between consecutive images I_t, are needed to solve for the flow values. Each frame is convolved in x, y, and t with a 1 x 5 kernel to compute I_x, I_y, and I_t, for each pixel. The kernel used in the implementation is shown in Figure 4.5 along with calculations of I_x and I_y using a sample set of image values. Gradients in x and y can be computed within the current frame, however I_t must be computed using consecutive temporal images.

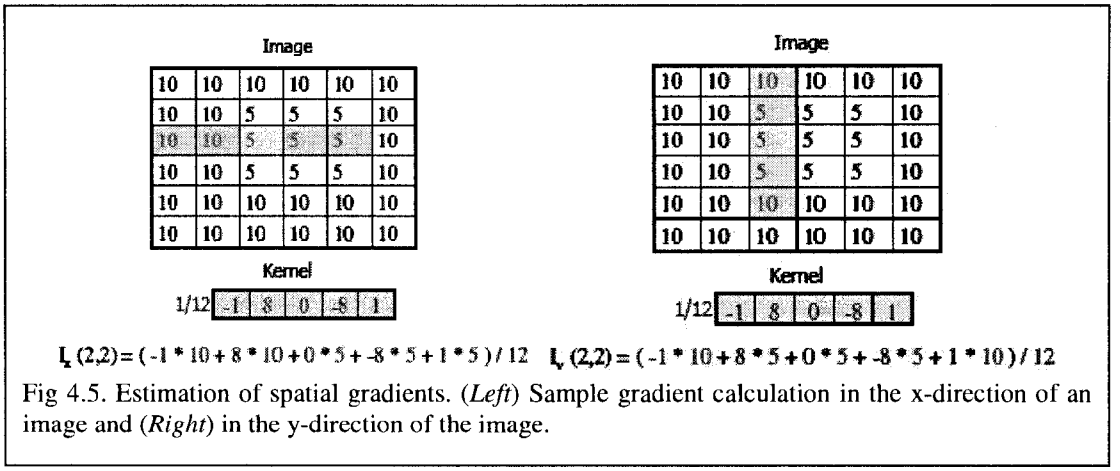


Fig 4.5. Estimation of spatial gradients. (*Left*) Sample gradient calculation in the x-direction of an image and (*Right*) in the y-direction of the image.

The solution obtained (u_{flow}, v_{flow}) is valid for objects experiencing small deformations. A modification proposed to remedy this problem is known as multi-resolution optical flow [125]. This consists of the following three steps: a) construction of Gaussian pyramids for each image, b) computation of optical flow at a coarse scale, and c) propagation of that optical flow to the next level of the pyramid. The first step is to create a Gaussian pyramid containing images at different scales. Each level of the pyramid should contain images scaled by a factor of two (and anti-aliased by virtue of the Gaussian filtering). Iteration over each level in the pyramid follows, and the optical flow is updated at each step. Once the optical flow is computed at a given resolution, it can be propagated to the next resolution by interpolation (Figure 4.6).

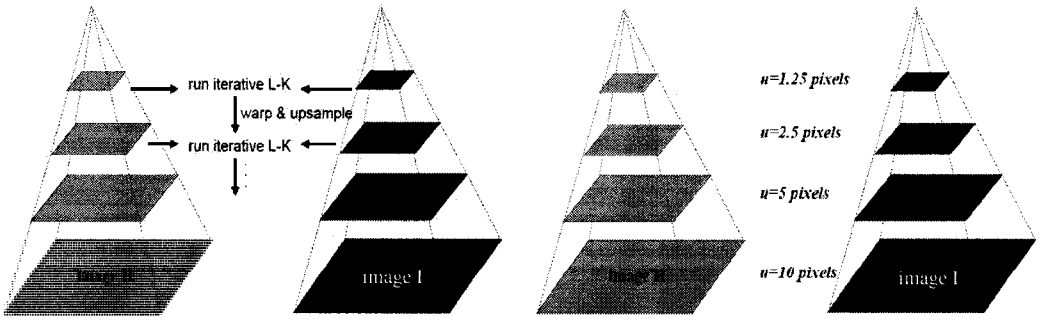


Fig 4.6. Coarse to fine optical flow estimation. (*Left pair*) The iterative refinement process begins with the estimation of the image velocities at each pixel by solving Lucas-Kanade equations. Then, image H is warped towards image I using the estimated flow field by using conventional image warping techniques. (*Right pair*) Ideal values obtained for deformations of about 10 pixels in the x pixel components. Image taken from [125].

The flow values also need to be scaled to account for resizing. The objective is to find the corresponding pixel location in a second image, *I*, from a pixel position in the first image, *H*. We achieve this by using the pyramidal approach and calculating the final optical flow vector, ***d***, yielding *H*= *I* + ***d***. The complete Pseudocode is presented in Figure 4.7.

Goal: Let $\mathbf{u}$ be a point on image I. Find its corresponding location $\mathbf{v}$ on image J

Build pyramid representations of I *and* J: $\{I^L\}_{L=0,\ldots,L_m}$ and $\{J^L\}_{L=0,\ldots,L_m}$

Initialization of pyramidal guess: $\mathbf{g}^{L_m} = [g_x^{L_m} \;\; g_x^{L_m}]^T = [0 \;\; 0]^T$

for $L = L_m$ **down to** 0 **with step of** -1

Location of point $\mathbf{u}$ *on image* I^L: $\mathbf{u}^L = [p_x \;\; p_y]^T = \mathbf{u}/2^L$

Derivative of I^L *with respect to* x: $I_x(x,y) = \dfrac{I^L(x+1,y) - I^L(x-1,y)}{2}$

Derivative of I^L *with respect to* y: $I_y(x,y) = \dfrac{I^L(x,y+1) - I^L(x,y-1)}{2}$

Spatial gradient matrix: $G = \displaystyle\sum_{x=p_x-\omega_x}^{p_x+\omega_x} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} \begin{bmatrix} I_x^2(x,y) & I_x(x,y)I_y(x,y) \\ I_x(x,y)I_y(x,y) & I_y^2(x,y) \end{bmatrix}$

Initialization of iterative L-K: $\overline{v}^0 = [0 \;\; 0]^T$

for $k = 1$ **to** K **with step of** 1 (or until $\|\overline{\eta}^k\| <$ accuracy threshold)

Image difference: $\delta I_k(x,y) = I^L(x,y) - J^L(x + g_x^L + v_x^{k-1}, y + g_y^L + v_y^{k-1})$

Image mismatch vector: $\overline{b}_k = \displaystyle\sum_{x=p_x-\omega_x}^{p_x+\omega_x} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} \begin{bmatrix} \delta I_k(x,y) I_x(x,y) \\ \delta I_k(x,y) I_y(x,y) \end{bmatrix}$

Optical flow (Lucas-Kanade): $\overline{\eta}^k = G^{-1}\,\overline{b}_k$

Guess for next iteration: $\overline{v}^k = \overline{v}^{k-1} + \overline{\eta}^k$

end of for-loop on k

Final optical flow at level L: $\mathbf{d}^L = \overline{v}^K$

Guess for next level $L-1$: $\mathbf{g}^{L-1} = [g_x^{L-1} \;\; g_y^{L-1}]^T = 2\,(\mathbf{g}^L + \mathbf{d}^L)$

end of for-loop on L

Final optical flow vector: $\mathbf{d} = \mathbf{g}^0 + \mathbf{d}^0$

Location of point on J: $\mathbf{v} = \mathbf{u} + \mathbf{d}$

Solution: The corresponding point is at location $\mathbf{v}$ on image J

Fig 4.7. Pseudocode of the displacement estimation of structures between images using a pyramid optical flow approach [125].

4.4.2 Gradient Vector Flow & Snakes

Active contour models have been vastly used in many tracking and segmentation problems and they are represented by a parametric curve x = [*x(s), y(s)*] that deforms through the image domain to minimize an energy function. They minimize the following Euler equation:

$$\alpha x'' - \beta x'''' - \nabla E_{ext} = 0 \tag{4.12}$$

The first two parameters α and β are weights explicitly used for the internal bending energy and ∇ is a gradient operator. The gradient vector flow field (GVF) is not formulated using the standard energy minimization framework, but is constructed by a force balance condition. As developed in [126], the GVF is a vector field that minimizes the energy function:

$$E = \iint (\mu(|\nabla u|^2 + |\nabla v|^2) + |\nabla f|^2 |v - \nabla f|^2) dxdy \tag{4.13}$$

where (u, v) is the vector field and f is the edge map of the input image with μ being a parameter controlling the smoothness of the GVF field. Traditional edge maps are computed using Gaussian convolutions at a specified scale, σ. The external energy term in (4.12) is replaced by (4.13) to yield the final active contour formulation:

$$\alpha x'' - \beta x'''' + v = 0 \tag{4.14}$$

where x is the active contour curve. Lastly, we introduce a third and final parameter, τ, which is defined as the iteration time step. By normalizing the external force, each contour point will move at most τ pixels per iteration (example: 0.1 pixel time steps). Figure 4.8 illustrates the major advantage of the gradient vector flow formulation. The active contour has a larger capture range for it to converge properly towards the targeted structure.

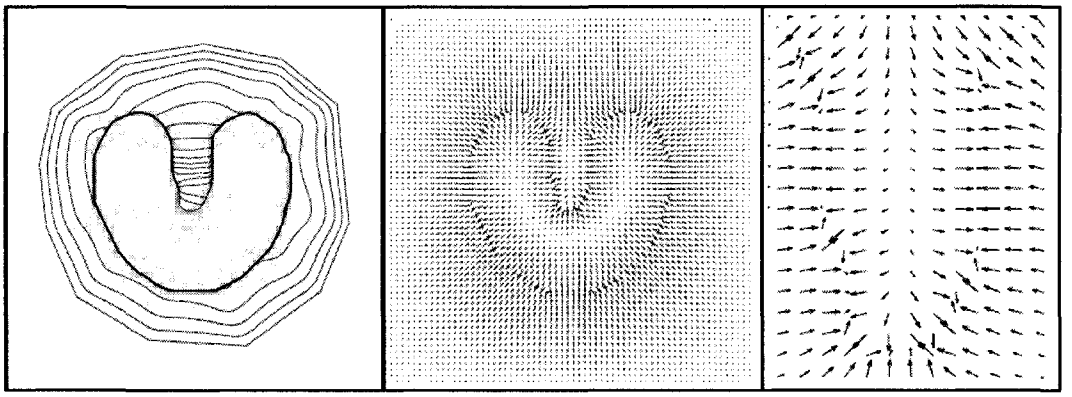


Fig 4.8. Parametric deformable model as derived in [109]. (*Left-Center*) The GVF external forces are derived from a diffusion operation, and they tend to extend very far away from the object. This extends the "capture range" so that snakes can find objects that are quite far away from the snake's initial position. (*Right*) This same diffusion creates forces which can pull active contours into concave regions.

The concavity capture range is imperative when dealing with coronary arteries that have a stenosis. The inherent U-shape of the stenosis reaffirms the necessity of implementing a GVF formulation for the segmentation and tracking procedures. The optical flow deformation values (u_{flow}, v_{flow}) will "boost" our centerline deformable contour much closer to the true position of the artery in the subsequent frame, in hopes of decreasing the capture range of the snake and having it successfully converge during the optimization step.

4.5 Camera Coordinate System and Projection Matrices

Figure 4.9 shows the full perspective camera model that will be used for the 3D reconstruction problem. If we define a three dimensional point P_{world}= [X Y Z 1]^T in the world coordinate system, then its 2D projection in an image, m= [u v 1]^T, is achieved by constructing a projection matrix:

$$P_{mat} = \begin{bmatrix} kf & 0 & u_o \\ 0 & kf & v_o \\ 0 & 0 & 1 \end{bmatrix} \times \left[\begin{array}{ccc|c} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{array}\right]$$

$$\lambda m = P_{mat} P_{world} \tag{4.15}$$

The intrinsic matrix of size [3x3], contains pixel coordinates of the image center, also known as the principal points (u_o, v_o), the scaling factor k, which defines the number of pixels per unit distance in image coordinates, and the focal length f of the camera (in meters). The extrinsic matrix of size [3x4] is identified by the transformation needed to align the world coordinate system to the camera coordinate system. This means that a translation vector, t, and a rotation matrix, R, needs to be found between the positions of the origins and in order to align the corresponding axis of the two reference frames respectively.

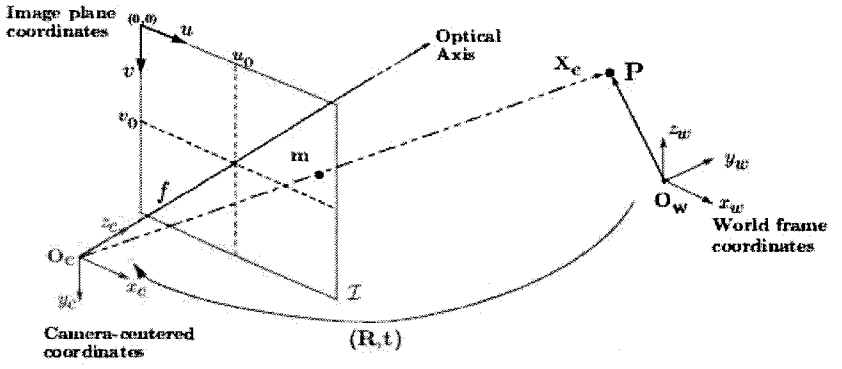


Fig 4.9. The perspective camera model. Any 3D world point can be projected onto a 2D plane and its coordinates would be (u, v) pixels. The camera model is taken from the Epipolar Geometry Toolbox [127].

4.5.1 Orthographic and Weak Perspective Camera Models

An orthographic camera is one that uses parallel projection to generate a two dimensional image of a three dimensional object. The image plane is perpendicular to the viewing direction. Parallel projections are less realistic than full perspective projections, however they have the advantage that parallel lines remain parallel in the projection, and distances are not distorted by perspective foreshortening. The orthographic projection matrix is given by

$$P_{ortho} = \begin{bmatrix} k*r_{11} & k*r_{12} & k*r_{13} & (k*t_x)+u_o \\ k*r_{21} & k*r_{22} & k*r_{23} & (k*t_y)+v_o \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.16}$$

The weak perspective camera is approximately the full perspective camera with individual depth points Z_i replaced by an average depth Z_{avg}. We define the average depth, Z_{avg} as being located at the centroid of the cloud of 3D points in the world coordinate system. The weak perspective projection matrix is given by

$$P_{weak} = \begin{bmatrix} f*k*r_{11} & f*k*r_{12} & f*k*r_{13} & (f*k*t_x)+(u_o*Z_{avg}) \\ f*k*r_{21} & f*k*r_{22} & f*k*r_{23} & (f*k*t_y)+(v_o*Z_{avg}) \\ 0 & 0 & 0 & Z_{avg} \end{bmatrix}$$

$$\text{where} \quad Z_{avg} = ([r_{31}\ r_{32}\ r_{33}]^T \times centroid) + t_z \tag{4.17}$$

Derivations are omitted here for the sake of brevity. Figure 4.10 shows a depiction of the approximate camera models.

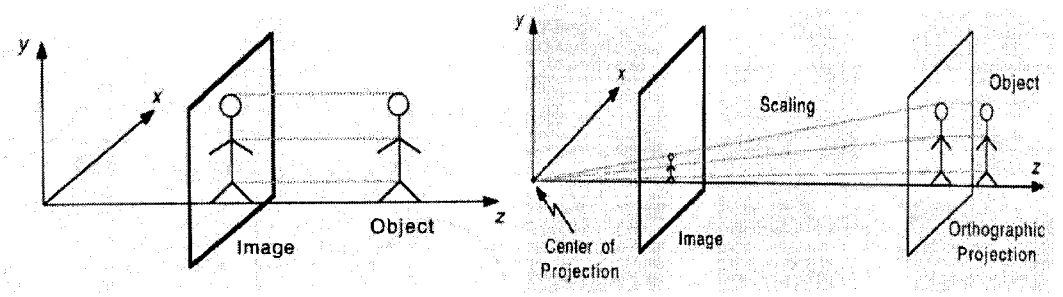


Fig 4.10. (*Left*) Orthographic camera model. It is described by a projection of a 3D object onto a plane by a set of parallel rays orthogonal to the image plane. (*Right*) Weak Perspective camera model. It is equivalent to a scaled orthographic projection where the projection of each 3-D point is not affected by their respective depth coordinates but rather by the average scene depth.

In the clinical setting, the monoplane C-arm fluoroscope parameters are usually given in a header file that comes included with the angiographic datasets (Figure 4.11). Thus, the focal length described above would be equivalent to the SID variable of the C-arm, and the rotation matrices above would need to be calculated using the PA and SA values.

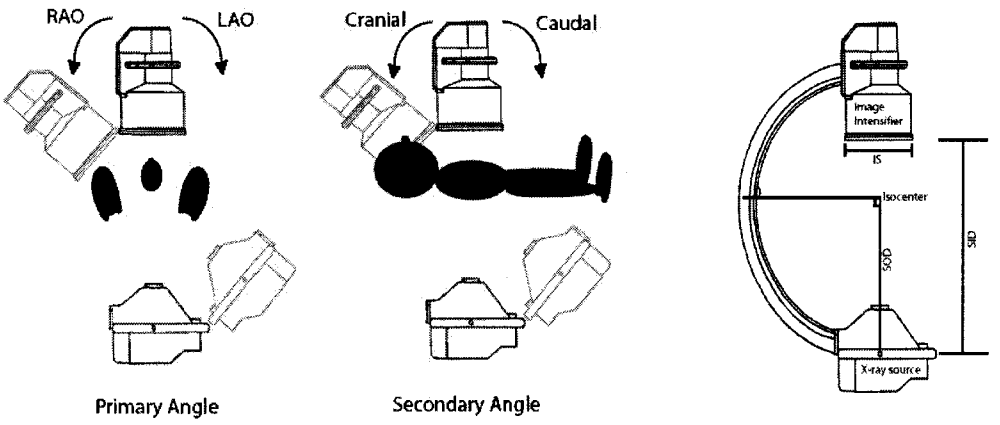


Fig 4.11. Gantry parameters for a typical monoplane C-arm fluoroscope. (*Left*) The rotation angles with respect to a patient can be decomposed to primary and secondary angle of rotations. (*Right*) the intensifier size determines the size of pixels with respect to world coordinates, and the source to object distance will be defined as half the value of the source to intensifier distance.

4.6 Optimal Correspondence for Biplane 3D Reconstruction

4.6.1 3D Biplane Reconstruction

The biplane algorithm used is a standard algorithm where both projection matrices are *known a priori*, due to the availability of the gantry parameters from the dataset header files for both images, and the 3D world points are determined by solving an over determined set of equations, as in the following:

$$u_1(\xi_1) = \frac{P^1_{mat1}(\xi_1)\times(X\ Y\ Z\ 1)^T}{P^3_{mat1}(\xi_1)\times(X\ Y\ Z\ 1)^T} \qquad v_1(\xi_1) = \frac{P^2_{mat1}(\xi_1)\times(X\ Y\ Z\ 1)^T}{P^3_{mat1}(\xi_1)\times(X\ Y\ Z\ 1)^T}$$

$$u_2(\xi_2) = \frac{P^1_{mat2}(\xi_1)\times(X\ Y\ Z\ 1)^T}{P^3_{mat2}(\xi_1)\times(X\ Y\ Z\ 1)^T} \qquad v_2(\xi_2) = \frac{P^2_{mat2}(\xi_1)\times(X\ Y\ Z\ 1)^T}{P^3_{mat2}(\xi_1)\times(X\ Y\ Z\ 1)^T} \tag{4.18}$$

The above four equations with three unknowns (X, Y, Z) are governed by the 6 gantry parameters ξ_i= [*SID, SOD, PA, SA,* u_o, v_o] for i=1, 2 image views. Further, P^j_{mat}, for j=1, 2, 3 signifies the j^{th} line of the projection matrix.

4.6.2 RANSAC & Geometric Curvature Constraint

As stated in the previous chapter, a major problem encountered when performing 3D reconstruction using multiple view algorithms is the matching of the feature points between the images.

Recalling that the fundamental matrix, F, is the algebraic representation of the epipolar geometry, its precision is affected by outlier errors, and robust methods need to be employed when calculating it. RANSAC estimates F by randomly selecting the minimum required number of correspondences. Given a number of points k needed to estimate the RANSAC model, S subsets of samples of k points are first extracted from

the data and used to estimate an initial set of S hypotheses. If the minimal number of data points k needed to estimate the model and the fraction of the outliers ε, are known, we can compute the confidence ρ that given s samples, at least one of the samples is outlier free as:

$$\rho = 1-(1-(1-\varepsilon)^k)^s \tag{4.19}$$

To achieve a desired confidence ρ we can compute from the above equation the required number of samples as follows:

$$s = \left\lceil \frac{\ln(1-\rho)}{\ln(1-(1-\varepsilon)^k)} \right\rceil \tag{4.20}$$

For each F computed, the set of inliers is calculated. The fundamental matrix with the highest consensus of inliers is selected to be the final computed F. This computation is repeated iteratively until a certain threshold is achieved. This threshold is equivalent to the inlier pixel distance with respect to the epipolar line. It is equal to the Sampson distance (geometric error) of the fit of a fundamental matrix with respect to a subset of corresponding points:

$$Sampson\ error = \frac{(x_i^T F_j x_i')^2}{(F_j x_i)_1^2 + (F_j x_i)_2^2 + (F_j^T x_i')_1^2 + (F_j^T x_i')_2^2} \tag{4.21}$$

As RANSAC is a robust method used primarily in the 3D vision environment where rigid motion of objects is predominant, we cannot rely on its solutions to obtain a final set of matching correspondences. Hence, we will rely on a method that will filter out the RANSAC optimum matches by taking into consideration the 2D information of the points, as well as their 3D geometrical reprojected curvature.

Li et al. [128] relate the 3D differential geometry of curves in space to the 2D differential geometry of the planar curves obtained on 2D images. This relationship can then be used to solve the stereo-correspondence problem. The local behavior of a 3D

curve can be determined by a third-order Taylor expansion given that we know the tangent, normal, binormal, curvature and torsion (T, N, B, κ, τ) at that point. In projective geometry, a 3D point will project onto the left image and right images of a stereo image pair. We define θ as the orientation of a tangent to a 2D point in the image. Corresponding 2D curve points are represented by their pixel positions, tangent orientation and curvature, that is, (x_l, y_l, θ_l, κ_l) and (x_r, y_r, θ_r, κ_r) in the left and right image respectively. These are called image corresponding pairs. Li shows that by using projective properties of a unit sphere centered at the camera, that the 3D normal N and the curvature κ can be recovered. Hence, given image corresponding pairs taken from specific X-ray gantry settings, the 3D position X and the 3D tangent T can be computed using standard methods (epipolar geometry and Frenet-Serret equations respectively). Then in order to solve for the 3D normal, N, and curvature, κ, we need to know the normalized vector pointing towards the projection, p, the focal length f that is assumed to be the same for both cameras, the distance between 3D point and camera center, λ, and the 2D tangent and curvature in the image (Figure 4.12). The following equations are used to relate 2D/3D curvature.

$$
\begin{aligned}
(p_l \times T)\cdot N\kappa &= \frac{f(1-(p_l\cdot T)^2)^{3/2}}{\lambda_l(1-(p_l\cdot t_l)^2)^{3/2}}\kappa_l \quad (left\,image)\\
(p_r \times T)\cdot N\kappa &= \frac{f(1-(p_r\cdot T)^2)^{3/2}}{\lambda_r(1-(p_r\cdot t_r)^2)^{3/2}}\kappa_r \quad (right\,image)\\
T\cdot N\kappa &= 0
\end{aligned}
\tag{4.22}
$$

This is a linear system with four unknowns subject to the constraint that $\|N\| = 1$, which will be treated as three unknowns $N\kappa = [N_x\kappa, N_y\kappa, N_z\kappa]$. The optimized three dimensional curvature, κ, can be solved for as it is equal to $\|N\kappa\|$. Corresponding points in images that reproject into the 3D world coordinate system and that have a κ similar to the 3D Frenet-Serret curvature values of the coronary arteries will be retained as potential candidates.

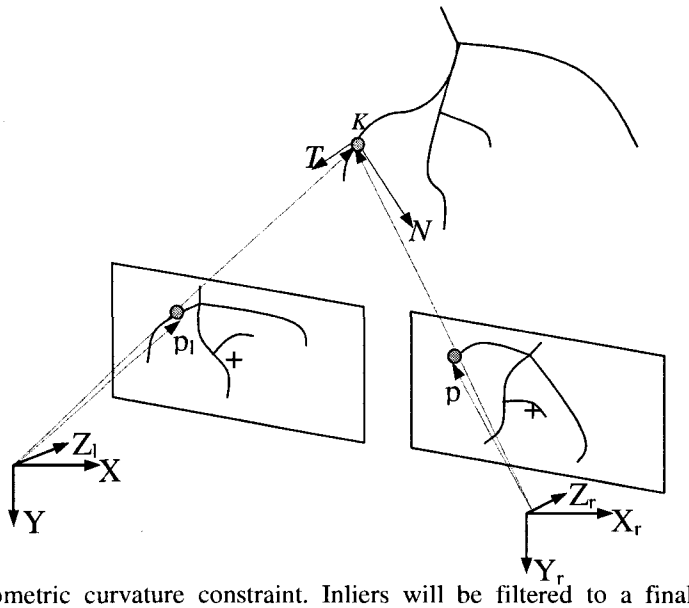


Fig 4.12. Geometric curvature constraint. Inliers will be filtered to a final subset by considering that 2D image correspondences between the left and right images must have similar curvature when back projected into the 3D world space. The normal N and curvature κ in three dimensions can be calculated if estimated 3D points and tangents T are available a priori, as well as, the 2D Frenet information (tangent and curvature) for 2D correspondences.

The derivation of the above constraints is as follows. We first assume that a smooth 3D curve with non-vanishing curvature is constructed in three dimensions. We denote its position by the vector r(s) in the world coordinate system, and its spherical projection to the unit radius image sphere with center **c** by a vector **u** in the camera coordinate system (Figure 4.13).

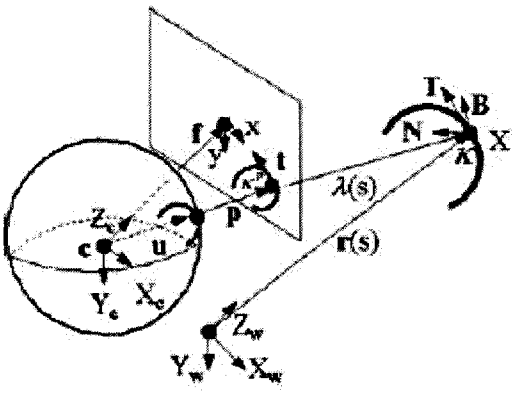


Fig 4.13. 3D space curve, its spherical projection to unit sphere, and its perspective projection to image plane. Image taken from [128].

Assume the world and the camera coordinate systems have the same orientation, that is, there is only translation (no rotations between them. Then the space curve can be described as: $\mathbf{r}(s) = \mathbf{c} + \lambda(s)\mathbf{u}(s)$, where $\mathbf{c}$ is the vector pointing to the camera center in the world coordinate system, λ is the Euclidean distance of $\mathbf{r}(s)$ to the camera coordinate system origin, and $\mathbf{u}$ is the vector pointing to its spherical projection in the camera coordinate system.

Previous work [129, 130] shows that the relationship of image geodesic curvature and the 3D space curve geometry under perspective projection is:

$$\kappa^g = \frac{\lambda\kappa(u\times T)\cdot N}{(1-(u\cdot T)^2)^{3/2}} \tag{4.23}$$

where κ^g is the image geodesic curvature at the image spherical projection $\mathbf{u}$ of the space point being studied, T and N are the space tangent and normal at that point, κ is the curvature at the space point $\mathbf{r}(s)$. We now calculate the relationship between 3D space curve properties and their (measurable) perspective projection image plane properties.

Consider the 2D image plane curve formed by $\mathbf{r}(s)$ through perspective projection, with $\mathbf{p}$ the vector pointing to the projection in the image plane ($\mathbf{u} = \mathbf{p}/\|\mathbf{p}\|$), and $\mathbf{f}$ the vector pointing to the image center, both in the camera coordinate system.

Since we choose the camera and image coordinate systems such that they are aligned, two dimensional vectors tangent and normal in the image coordinate system can be lifted to R^3, the camera coordinate system, to **t** and **n**, with the third component being zero. Note that this construction will form the same spherical projection as **r**(s). Furthermore, although for a plane curve the curvature is defined as the signed curvature κ^p, and the normal is defined such that the tangent-normal basis has the same orientation as the coordinate axis basis, the quantity $\kappa^p\mathbf{n}$ is still the same if we study this curve as a 3D curve. The binormal at any point on this curve is $\mathbf{b} = \mathbf{t} \times \mathbf{n} = \mathbf{f}/\|\mathbf{f}\|$. The relationship between the perspective projection image curvature and the geodesic curvature is thus given by:

$$\begin{aligned}\kappa^g &= \frac{\|p\|\kappa^p (u\times t)\cdot n}{(1-(u\cdot t)^2)^{3/2}} = \frac{\|p\|\kappa^p (u\cdot(t\times n)}{(1-(u\cdot t)^2)^{3/2}} \\ &= \frac{\|p\|\kappa^p u\cdot\frac{f}{\|f\|}}{(1-(u\cdot t)^2)^{3/2}} = \frac{\frac{1}{\|f\|}\kappa^p p\cdot f}{(1-(u\cdot t)^2)^{3/2}} \\ &= \frac{\|f\|\kappa^p}{(1-(u\cdot t)^2)^{3/2}} = \frac{f\kappa^p}{(1-(u\cdot t)^2)^{3/2}}\end{aligned} \tag{4.24}$$

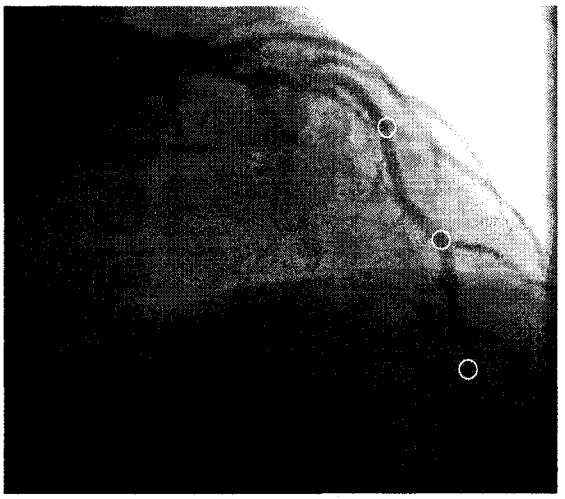
with f=$\|\mathbf{f}\|$ the focal length obtained from the camera calibration. Using the geodesic curvature as the bridge, we can now connect the above two equations to formulate the relationship of the 3D space curve curvature and its perspective projection image curvature:

$$\frac{\lambda(u\times T)\cdot N}{(1-(u\cdot T)^2)^{3/2}}\kappa = \frac{f}{(1-(u\cdot t)^2)^{3/2}}\kappa^p \tag{4.25}$$

This final equation gives the first two constraint equations for both views in the non simultaneous image setup.

4.7 Three Dimensional Reconstruction using Monoplane Angiographic Images

In order to test the feasibility of reconstructing the coronary arteries in three dimensions using only a single view, we need to take advantage of the information readily available to us. First, we know that the X-ray gantry parameter settings remain unchanged for a given angiographic dataset. This is important as we can create a projection matrix from the given parameter values in the header file. An example of such values are illustrated in (Figure 4.14-right).



Distance Source to Detector (mm)	"1080"
Intensifier Size (mm)	"178"
Imager Pixel Spacing (mm/pixel)	"0.3704 "
Positioner Primary Angle (degrees)	"-34.0 "
Positioner Secondary Angle (degrees)	" 30 "
Rows (pixels)	"512"
Columns (pixels)	"512"
Window Center (pixels)	"128 "
Window Width (pixels)	"255 "

Fig 4.14. (*Left*) Monoplane coronary angiogram. Regions of potential rigid motion, such as the coronary bifurcation points, are shown circled. (*Right*) The self calibration of a monoplane C-arm fluoroscope involves estimating the gantry parameter values.

Secondly, we can make use of pre-operative volume information by either making use of computed tomography (CT), magnetic resonance imaging (MRI), or by using rotational X-ray images (3 or more projections at different angulations). In the CT and MRI case, by applying manual segmentation or advanced segmentation techniques such as level sets for example, we can obtain a 3D volume of the coronary arteries that are being investigated. On the other hand, if we are provided with images from a rotational X-ray system, then applying a three view reconstruction algorithm can yield a suitable 3D

volume. Lastly, we see no disadvantage of reconstructing a cardiac structure using the biplane reconstruction algorithm. In fact, it may be advantageous pursuing this path as the X-ray fluoroscope is already located in the intervention room, hence there is no need to rely on pre-operative data coming from another modality technology.

The strategy of our proposed monoplane reconstruction algorithm is as follows. Our first hypothesis is that we can exploit the rigid movements experienced by certain regions of the coronary arteries as they deform through the cardiac cycle. As an example of such rigid regions, the bifurcation points may be considered or we can concentrate on a section of the artery where a stent may have been implanted to restore blood flow. These regions technically experience movements of a non-contractive nature when compared to other parts of the coronary arteries (Figure 4.14-left).

Therefore, if we suppose that we have at our disposal a set of n 3D points ($X0_n$, $Y0_n$, $Z0_n$) at time $t = 0$ with the use of the pre-operative data, then we are able to optimize the gantry parameters of the X-ray system by using a standard direct linear transformation (DLT) algorithm. Figure 4.15 demonstrates the 2D/3D registration procedure between the two different modalities. As the registration scope is beyond the thesis objectives, we focus on implementing the DLT calibration to obtain the projection matrix. Having optimized the gantry system parameters we can now solve for the 3D displacements (dx_n, dy_n, dz_n) between subsequent X-ray image frames.

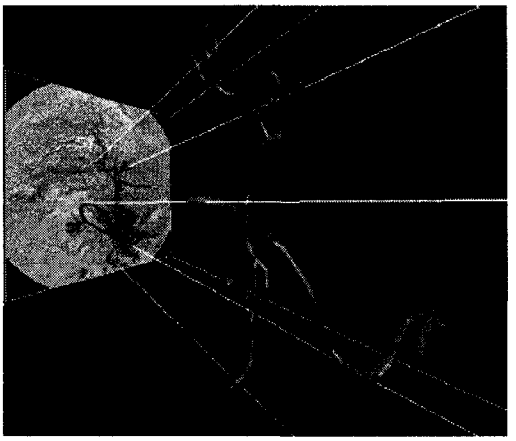


Fig 4.15. Visualization of the 2D/3D scenario. The 3D vasculature obtained from either MRI or CT scans is to be deformably registered to the 2D X-ray projection image. Image taken from [131].

4.7.1 Minimum of Three Images for Monoplane 3D reconstruction

Expanding equations (4.18) and using an additional angiographic image at time $t = 1$, we obtain our first two equations as follows,

$$u_1 = \frac{m_1(X0+dx_1)+m_2(Y0+dy_1)+m_3(Z0+dz_1)+m_4}{m_9(X0+dx_1)+m_{10}(Y0+dy_1)+m_{11}(Z0+dz_1)+m_{12}}$$
$$v_1 = \frac{m_5(X0+dx_1)+m_6(Y0+dy_1)+m_7(Z0+dz_1)+m_8}{m_9(X0+dx_1)+m_{10}(Y0+dy_1)+m_{11}(Z0+dz_1)+m_{12}} \quad (4.26)$$

Both equations describe the pixel coordinates in an image (u, v) and the twelve coefficients $m_{i=1:11}$ are the values of the projection matrix components. By adding an additional X-ray image frame at time $t = 2$ we obtain two new equations with three additional unknowns in 3D

$$u_2 = \frac{m_1(X0+dx_1+dx_2)+m_2(Y0+dy_1+dy_2)+m_3(Z0+dz_1+dz_2)+m_4}{m_9(X0+dx_1+dx_2)+m_{10}(Y0+dy_1+dy_2)+m_{11}(Z0+dz_1+dz_2)+m_{12}}$$

$$v_2 = \frac{m_5(X0+dx_1+dx_2)+m_6(Y0+dy_1+dy_2)+m_7(Z0+dz_1+dz_2)+m_8}{m_9(X0+dx_1+dx_2)+m_{10}(Y0+dy_1+dy_2)+m_{11}(Z0+dz_1+dz_2)+m_{12}} \quad (4.27)$$

These first four equations take into account the spatial positions of a projected 3D world point on the X-ray images. We can extract two additional equations on the basis that we know the euclidean distances, *d*, between image points in two consecutive images.

$$d_{01}{}^2 = \sqrt{(u_1-u_0)^2+(v_1-v_0)^2} \quad \text{using images at } t=0 \text{ and } t=1$$

$$d_{12}{}^2 = \sqrt{(u_2-u_1)^2+(v_2-v_1)^2} \quad \text{using images at t = 1 and t = 2} \quad (4.28)$$

The equations defined over a set of three angiographic images are the strict minimum in order to solve for the three dimensional displacements. A Levenberg-Marquardt optimization scheme [132] can be used here in order to solve for the unknown displacements. For the optimization scheme, initial approximations are a requirement to initialize the process. Hence, a suitable approximation for the displacements *dx* and *dy* can be obtained if we consider an orthographic back projection of the 2D image points. As for the displacement *dz*, two possible scenarios arise when back projecting an image point to the world reference frame (Figure 4.16). We expect that a mean displacement of the depth, using a weak perspective model, is a suitable approximation for *dz*. If we assume that the average depth of the coronary artery remains relatively constant in consecutive time frames, then we can readily calculate the average depth of the 3D points (*X0, Y0, Z0*) at time $t = 0$. The mean depth calculated will be equal at time instants $t = 1, 2$, etc., signifying that dz_1 and dz_2 will be set to zero for the optimization scheme. However, for the sake of a more exhaustive analysis, we will also consider depth displacements $dz \in [1, 2]$ millimeters as well. The reason for this is to provide

reasonable comparisons within a range of suitable approximations to validate the monoplane algorithm. For example, we deem depth displacements ranging between [5-10] mm unrealistic and thus settled for the three possible values we proposed as a logical starting point for our algorithm.

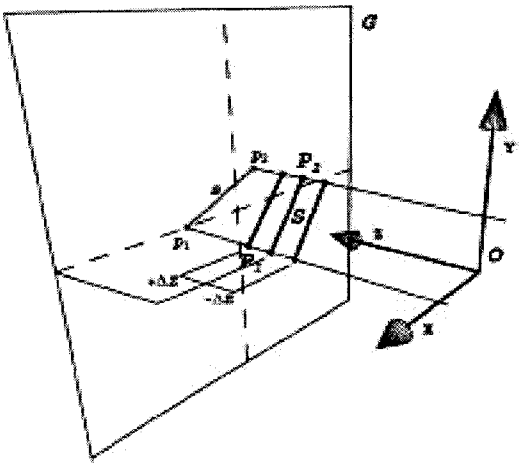


Fig 4.16. When considering backprojection of a 2D point to a 3D world point, two solutions might be possible when estimating the depth of a point. Here, we will assume an average depth of the coronaries, dz, and an orthographic projection approximation for dx and dy. Image taken from [118].

In the worst case, comparing the full perspective camera model to the weak perspective model gives the same angular variation as when an initial 3D point location is misestimated, using full perspective geometry alone. This can be a valid hypothesis when temporally tracking artery corresponding points that lie close to one another but would degenerate if we temporally tracked individual bifurcation points as they can lie very far from one another in the X-ray image.

We conclude by stating that the monoplane reconstruction algorithm proposed can be extrapolated by considering more than three consecutive images and that any other cardiac structures or objects (i.e. aorta, catheters) may be considered for reconstruction purposes.

4.8 Algorithm Validations

We present a brief overview of the validation process for each algorithm proposed previously. The choice for algorithmic parameters will be justified in each section of the following chapter.

4.8.1 Automatic Image Extraction Validation

The proposed extraction algorithm will be tested on six single view coronary datasets. As we did not have simultaneous electrocardiogram data to strengthen our validation, we simply looked at the image frame that displayed the smallest standard deviation obtained from our algorithm. The visual inspection should conclude that the image frame presents the least movement when compared to the subsequent images representing the cardiac cycle.

4.8.2 Angiographic Image Enhancement Validation

Our 4-step filter was applied to seven distinct angiographic datasets. Most of the parameters in our filter were left to the optimal values as designed by the respective authors who developed the homomorphic, anisotropic, and shock filters. The remaining parameters were empirically determined by running tests on the images in order to obtain images that enhanced the most prominent arteries present in the 2D images. The Lorenz and Frangi filters were also implemented as a comparison. For both these multiscale filters, we chose ten optimal scale values ranging between [0.25-2.5] pixels to provide the best chance in obtaining a high quality enhanced image in terms of background suppression and vessel enhancement.

4.8.3 Centerline Extraction Validation

The fast marching method was applied using our 4-step filtered image. This cost image served as a cost function to extract the centerline. To validate the accuracy of our two-click extracted centerline coordinates, we manually traced the medial axis on the targeted centerline and compared the manual 2D coordinates with those obtained by the FMM. The mean distance between the two coordinate sets was used for general quality assessment. Furthermore, we tested the extraction process on three different cost images. These three images were obtained by applying only: (i) the homomorphic filter, (ii) both the homomorphic and anisotropic filters and (iii) the homomorphic, anisotropic and shock filters. These cost images were used to verify the importance of applying all four filters to obtain a valid cost function in order to extract the artery centerlines.

4.8.4 Temporal Tracking Validation

Our tracking algorithm was tested on three different angiographic datasets on a total of 38 images. Two of the datasets showed a potential stenosis whereas the last dataset showed a healthy coronary tree. Validation was performed on both the pyramidal optical flow approach and the GVF external image fields. We compared the classical optical flow formulation to the pyramidal Lucas Kanade implementation by illustrating the estimated centerlines obtained in preceding images. The GVF fields were calculated using our 4-step filtered image and compared to the traditional edge images used in its formulation. Lastly, the active contour parameters were chosen to be the optimal ones provided by the authors who developed the GVF technique. Visual inspection was used to categorize a successful convergence of the active contour towards the centerline of the artery.

4.8.5 Optimal Point Correspondence Validation

Synthetic and clinical validation was performed to verify our novel curvature constraint. First, two traditional gantry setups were chosen when performing synthetic experimentation. These were the LAO/RAO and posterior/anterior viewing angles. We created a set of 3D coronary coordinates at a first time instant and deformed them through time and reprojected their points in a minimum of 5 sets of biplane images. We then obtained possible point matches using the RANSAC method. Lastly, we performed 3D reconstruction using those matches and the refined matches from the curvature constraint and compared the 3D RMS values. Clinical validation was performed using two distinct non simultaneous datasets by first self calibrating the gantry parameters using the curvature constraint point matches and then performing 3D reconstruction.

4.8.6 Monoplane 3D Reconstruction Validation

To validate our proposed monoplane algorithm we performed first rigid and non-rigid synthetic experimentation of structures. Since a centerline can be represented by a helix, we created a 3D helix in space and applied only rigid movement to it in a temporal fashion. The 3D points were reprojected in only single view 2D images. We also had at our disposal 13 ventricle markers of a sheep with their temporal 3D coordinates. These markers represent well the inherent non-rigid movement of the heart. We again reprojected these 3D markers on single plane synthetic images. Our monoplane algorithm was tested on the two types of movements.

Clinical validation was performed using 20 sets of biplane images showing ablation catheters in the left ventricle of a dog. One biplane image represented the left lateral view, whereas the second image represented the posterior/anterior view. The objective was to estimate the depth of the tip electrode of the ablation catheter using our proposed monoplane procedure. Therefore, for each biplane dataset, we first reconstructed the catheter in 3D using the two view algorithm, and then we used this a priori 3D reconstruction to estimate the 3D geometry on one of the views above. The

estimation of the 3D geometry was performed using three consecutive X-ray images and five consecutive X-ray images. Results were presented in terms of 3D RMS errors obtained between our monoplane algorithm and the reconstruction from the 20 biplane datasets.

Chapter 5- Results & Discussion

In this chapter, we will present the results obtained from the outlined methodology. First, we will show results for the different tools required to meet the two research hypothesis. These include: (i) extraction of the diastolic cardiac phase image, (ii) angiographic image segmentation, (iii) centerline extraction and (iv) temporal tracking. Then, we detail results for both hypotheses. These include the biplane matching process as well as the monoplane reconstruction algorithm. A general discussion will follow at the end of the chapter that will include the limitations of our methods and results.

5.1 Automatic Image Extraction

We tested the results on six angiographic datasets obtained from the Royal Victoria Hospital. Each dataset had an acquisition frame rate of 30 frames per second. This was equivalent to about 2 seconds of the cardiac cycle. We extracted a total of 65 image frames for each dataset and then calculated the standard deviation of the difference between consecutive image frames. In Figure 5.1, we can visualize the results obtained using the automatic extraction algorithm. It is to note that two apparent phase curves are visible before and after the 1 second time instant. In theory, two minimums can possibly be determined, one for each cardiac phase. The minimum peak corresponded to the smallest standard deviation, and hence the smallest heart motion, and by visualizing the actual angiographic image frame and its neighborhood, we confirmed that the image corresponded to the diastolic phase.

We then investigated the quality of the images. Visual inspection of the diastolic images confirmed that the image also had the least motion blur. However, the method assumes that we can actually see the contrast enhanced coronary artery in the angiographic image displaying the diastolic cardiac phase. If this was not the case, the diastolic image frame was selected as the first standard deviation minimum located just

after the 1 second time instant, which is also the beginning of the second cardiac cycle (see 2nd image in second column of Figure 5.1). The main contribution of such a method is that it can potentially be used instead of acquiring simultaneous electrocardiograms (ECG) in order to extract the relevant images.

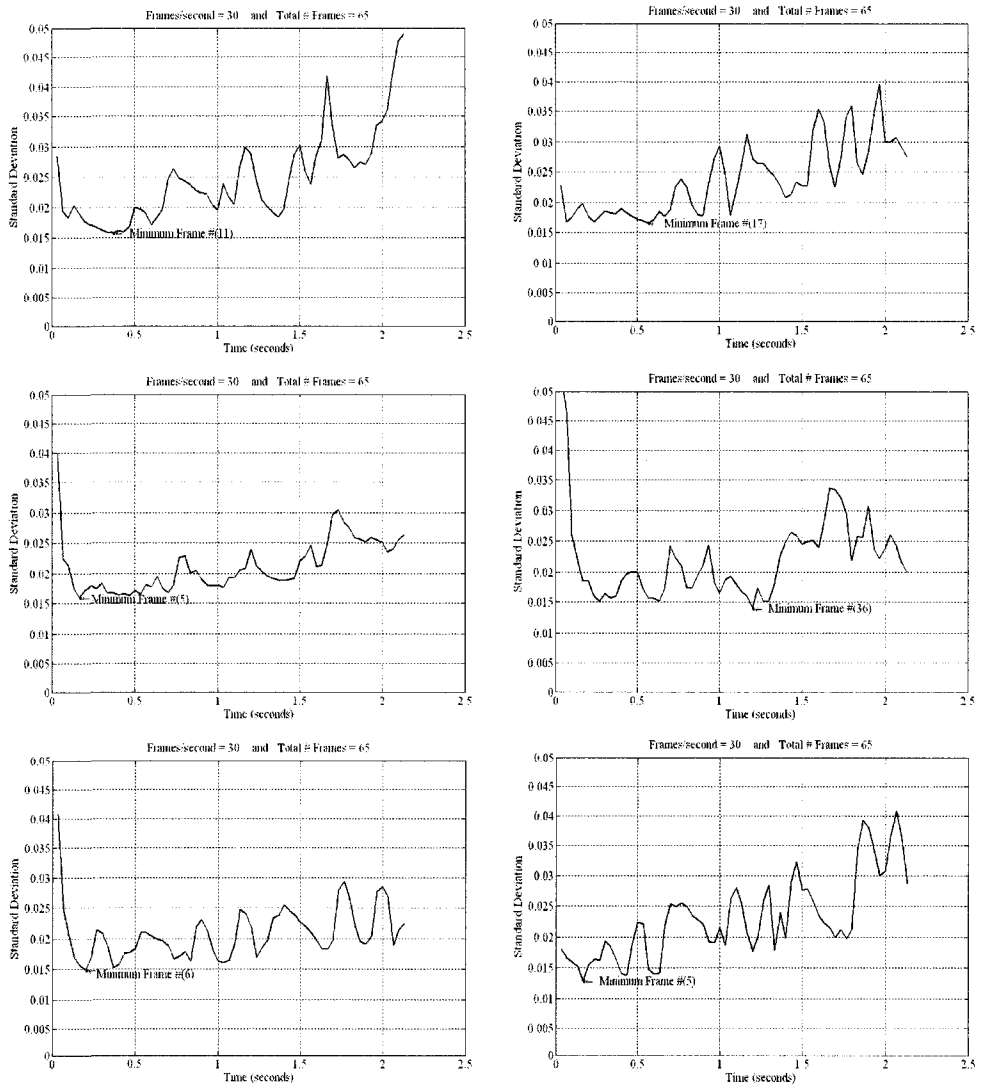


Fig 5.1. Automatic image extraction algorithm applied on six angiographic sequences. The minimum peaks (smallest standard deviation) in all figures corresponded to the diastolic phase.

5.2 Angiographic X-ray Image Segmentation

We added another angiographic dataset to bring the total to 7 different sets of images for image segmentation validation. The 4-step filter implementation first relies on a homomorphic filter to enhance contrast. This form of filter sharpens image features and flattens lighting variations in the X-ray image. The cutoff frequency of the filter was set to 0.5 and the ratio that high frequency values are boosted relative to the low frequency values was fixed to 2. These values were chosen so as to have a balance between visual contrasts and avoid losing contour information with respect to the principal coronary arteries. These two filter parameter values are to be unchanged when implementing the 4-step filter as they were determined empirically in order to arrive at a final image that will lead to positive results when extracting the artery centerline.

The Perona-Malik anisotropic diffusion filter was next implemented to enhance the quality of texture present in the image. It eliminates aliasing and introduces less blur thus preserving more detail. There are two parameters that need to be fixed. The parameter *kappa,* κ, is fixed to 25 as it controls conduction as a function of the image gradient. If κ is low then small intensity gradients are able to block conduction and hence diffusion across step edges (artery contours). A large value reduces the influence of intensity gradients on conduction. The parameter *lambda* controls speed of diffusion and we want it to be maximum so as to be time efficient. This parameter was set to 0.25. The number of iterations performed by the anisotropic filter was equal to 5. The parameter values, κ and *lambda,* were chosen to be the same ones as optimized by Perona-Malik when smoothing any input image.

The complex shock filter was then applied. The parameters chosen here were the same optimal ones as described by Gilboa et al. The complex scalar $\lambda = 0.1$, the real scalar was fixed to $\tilde{\lambda} = 0.2$, the parameter that controls the sharpness of the slope at the zero crossings of an artery edge is $a= 0.3$, and lastly the phase angle of complex part is fixed to $\theta=$ pi/1000. It was important to implement the complex shock filter because it not only smoothes texture perpendicular to the gradient, but also normal to it. Ideally,

the filter can be implemented at different phase angles in order to sweep the image in every direction in order to enhance all the arteries present in the X-ray image, regardless of their positions and angular configurations. The number of iterations for the complex shock filter was equal to 10. The shock filter parameters λ, $\tilde{\lambda}$, a, θ, were chosen to be the same as developed by Gilboa as they were optimized for noisy images such as angiography datasets. Had we increased the number of iterations drastically, the diffusion process would extend onto the coronary artery pixels and we would lose valuable perimeter information for the centerline extraction process. The same would be true when implementing the anisotropic diffusion process.

Lastly, the morphological operator had a structuring element the shape of a disk with a total diameter of 10 pixels, which is roughly the size of the diameter of the principal coronary arteries present in the X-ray images. In essence, the morphological filter will enhance objects of size 10 pixels in the image and suppress their background.

In Figure 5.2, we show the results of the Lorenz, Frangi and 4-step filter images applied to the four of the seven datasets. The response at different scales for the vessel enhancement filters can be combined by taking the maximum response over a range of scales. Hence, ten scales were used, $\sigma \in [0.25, 2.5]$ pixels. These were the scales that provided the best visual images as shown in Figure 5.2. Had we selected other ranges for the Gaussian scales, the Lorenz and Frangi images would have degraded in quality. This is a disadvantage of the vessel enhancement method. The main contribution of our method was to develop a filter that behaved in the same manner as the vessel enhancement filters- by suppressing the image background and enhancing the artery contours. Further, we wanted to implement a filter that had all of its parameters fixed for the user in hopes of diminishing user interaction for the selection of parameter processes. In clinical practice, we do not want to impose on the cardiologist any further time consuming steps such as "fine tuning" of algorithmic parameters when enhancing angiographic images. Ultimately, this was the goal we wanted to achieve in the end. The

results in Figure 5.2 show that our filter visually performed as well as the Lorenz and Frangi filters.

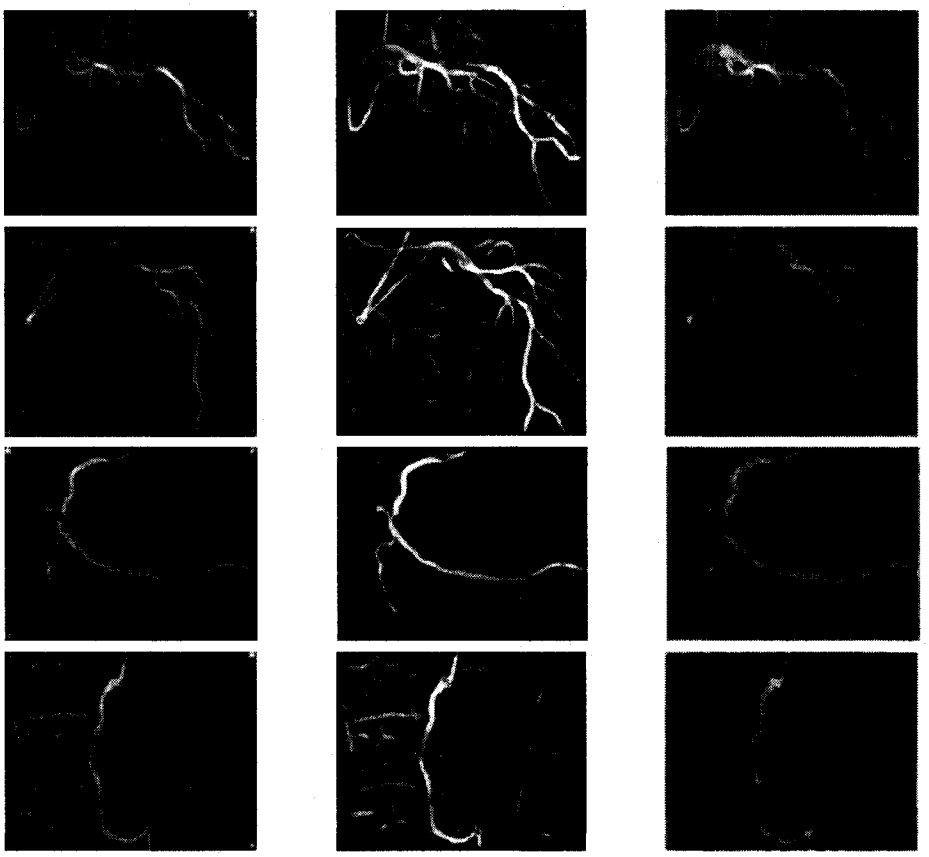


Fig 5.2. Vessel Enhancement Filters. (*Left column*) The Lorenz filter uses only the most predominant eigenvalues of the Hessian matrix in order to enhance the coronary arteries. (*Center column*) The Frangi filter uses all eigenvalues when enhancing the arteries. (*Right column*) Our proposed 4-step filter which is a combination of a homomorphic filter, anisotropic and shock filter and a morphological operator. The 4-step filter parameters remain constant, whereas the Gaussian scale used for the Lorenz and Frangi filters was set to 0.25 in this case. Had the Gaussian scale been larger in value, the filtered images would contain stronger artifacts.

5.3 Coronary Artery Centerline Extraction

The speed or cost function used to initialize the fast marching method (FMM) will be equivalent to the reciprocal of our vessel-enhanced image using our 4-step filter. It is the fourth step of the filter (morphological) that will permit the cost image to have a high cost for background pixels, and a low cost for coronary artery vessel pixels. To initialize the centerline extraction process, the user clicks two points equal to the principal coronary artery extremities. Results are evaluated based on an error measure corresponding to the distance *D(t)* between all points in the centerline path extracted by the fast marching method, *FMM(t),* and a reference path drawn by an observer representing the true centerline position of the artery, *Ref(s).* The error measure is computed as follows:

$$D(t) = \min_{s} \left| \mathrm{FMM}_{(t)} - \mathrm{Ref}_{(s)} \right| \tag{5.1}$$

The mean distance is used for general quality assessment. In Figure 5.3, some examples of coronary artery centerlines are shown on four of the seven datasets. The interventionist would select the coronary artery of interest to be investigated and then click its two extremities. The cost images worked satisfactorily considering the amount of noise present in the images.

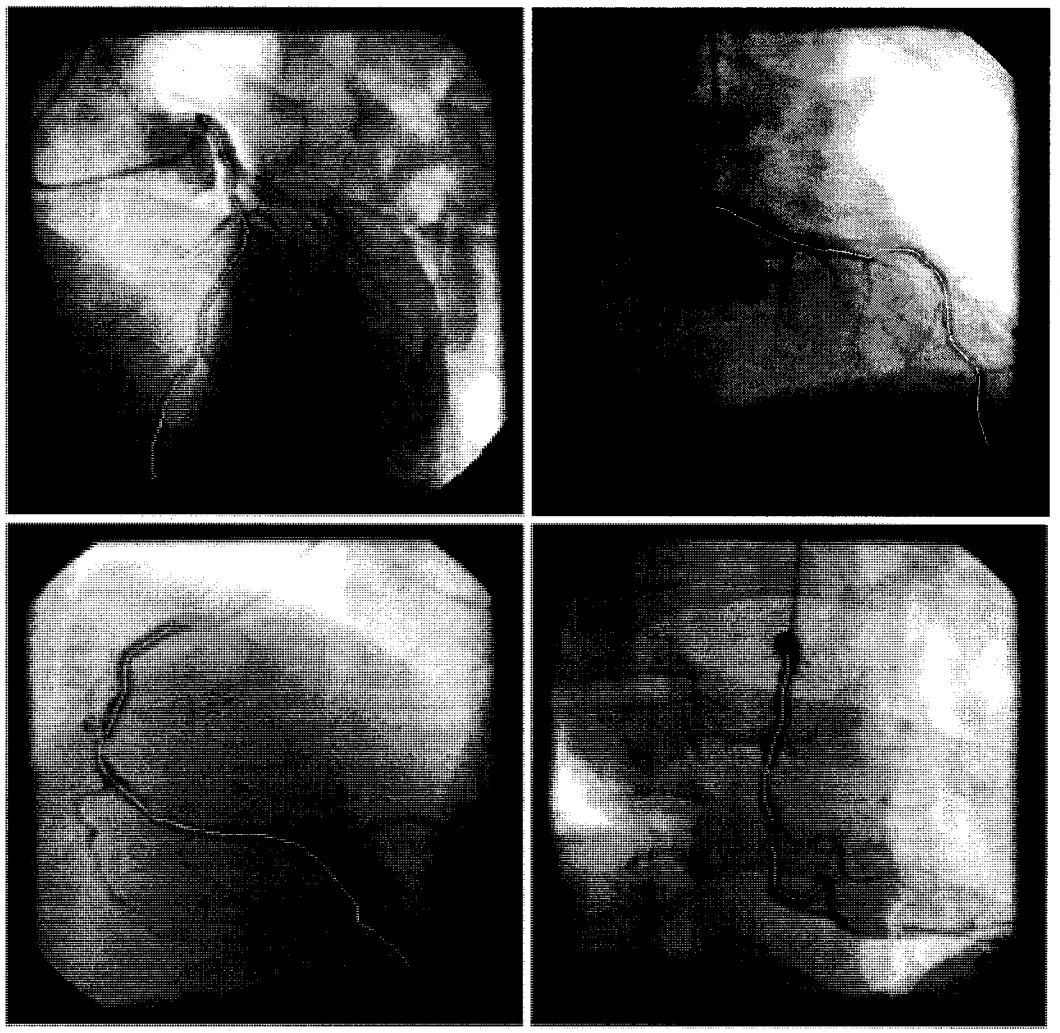


Fig 5.3. Centerline extraction results on angiographic images using our proposed 4-step filter and FMM method. The centerlines are superimposed on the original diastolic fluoroscopic images.

To justify the importance of the four preprocessing steps for our 4-step filter, we attempted to extract the centerlines using as a cost image: (i) only the homomorphic filtered image, (ii) the homomorphic and anisotropic filtered image, and (iii) the homomorphic filter coupled with shock and anisotropic diffusion image. In Figure 5.4, we note that the FMM was unable to cope with the cost images since an inaccurate path

was extracted between the extremities of the coronary artery. It is essential to include all four steps when using our empirically found parameter values.

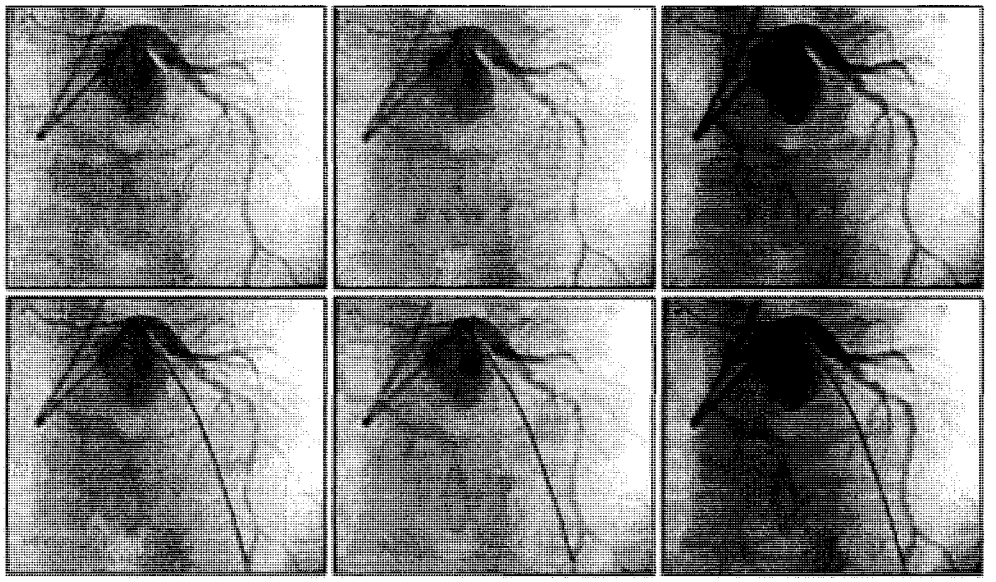


Fig 5.4. (*Top row*) Alternative cost images using only the homomorphic filter (*left*), both the homomorphic and anisotropic filters (*center*) and the homomorphic coupled with the shock and anisotropic diffusion (*right*). (*Bottom row*) Failure of the FMM using inadequate cost images (dotted line).

Lastly, Table 5.1 shows the results obtained in the study when comparing the automated centerline coordinates to those obtained manually. The Lorenz filter proved the more precise in 3 of the 7 images. The Frangi and 4-step filters had better average distance values in the remaining 4 images. Assuming a C- arm fluoroscope intensifier size of 178 mm, each pixel width was about 0.695mm. In all filters, the average distance errors were less than 0.58mm (that is the 0.83 pixel mean distance in the Lorenz filter). We were fortunate not to have many neighboring structures adjacent to our arteries during the vessel enhancement phase. If this were the case it might well lead to wrong trajectories in the FMM due to possible strong filter responses.

Table 5.1 Mean distances and STD's from FMM to Reference centerlines

	Lorenz Filter		Frangi Filter		4-Step Filter	
Image #	mean	std	mean	std	mean	std
1	0.5	0.31	0.42	0.29	0.58	0.41
2	0.83	0.54	0.81	0.52	0.62	0.4
3	0.37	0.25	0.31	0.25	0.56	0.27
4	0.48	0.27	0.63	0.43	0.7	0.6
5	0.64	0.83	0.77	0.76	0.55	0.34
6	0.41	0.3	0.44	0.36	0.57	0.41
7	0.37	0.34	0.45	0.3	0.6	0.46

The main contribution to our centerline extraction procedure is that the cost image to initialize the algorithm is obtained automatically via the 4-step filter. Although the interventionist is required to select two points to obtain the artery centerline, the method proved rigorous and robust on all seven datasets. Also, cost image was designed indirectly by the 4-step filter to take into account the stenosis positions, if present, on a given artery. The algorithm had no problem passing over a stenosis position, which sometimes appears as an empty hole or space on the artery perimeter.

5.4 Temporal tracking of the Coronary Arteries

As this was primarily a feasibility study, we had two datasets available where a stenosis was visible. Therefore it was decided to add a third dataset having no stenosis visible and perform the study with our proposed algorithm. Three angiographic datasets of different patients and 38 total image frames were used for evaluation of the tracking procedure. For each individual dataset, we extracted 15, 10, and 13 frames respectively that spanned one cardiac cycle. We first implemented our 4-step filter to the diastolic images and then extracted the principal coronary artery centerline using a two-click FMM approach. The centerline coordinates obtained in this first frame will be used to estimate the centerline positions in subsequent frames by using the multi-resolution optical flow method. A window region of 10 pixels was selected for the optical flow estimations to account for possible large deformations of the arteries, and 4 pyramid

levels were used for interpolation purposes to obtain a reliable flow estimate. Inside the windowed region we assume a constant flow hypothesis where all pixel values are assigned an equal weight of 1. For these centerlines to converge to their true positions, the GVF was implemented for all experiments using a regularization coefficient factor fixed at $\mu = 0.2$, as proposed by Xu et al., and the number of iterations equaled 15 so as to produce desired field vectors pointing toward the center of the artery (instead of pointing towards the artery wall contour, as in the traditional GVF formulation). The active contour parameters were empirically determined to be $\alpha = 1$, $\beta = 0.5$, and $\tau = 0.1$ as they yielded stable convergence of the coronary centerline throughout the total number of images tracked. In fact, we chose the same optimized parameter values as determined by Xu et al. when implementing their GVF active contour formulation.

The pyramidal optical flow approach took an average 14 minutes/frame, using MatLab, to obtain estimated centerlines when using two consecutive images. The mean displacement (± standard deviation) of the principal coronary artery between successive frames was: 7.57 (±2.44), 12.66 (± 6.42) and 10.54 (± 3.88) pixels. Figure 5.5 illustrates the difference between the traditional and pyramid Lucas Kanade optical flow approaches. In both (Figure 5.5a, b top rows), the traditional implementation for the optical flow fails to produce reliable centerline estimates due to large deformities of the artery. If deformations are larger than 10 pixels, the pyramidal optical flow produces reliable centerline position estimates but also shows signs of unreliability as seen by the elliptical region of interest in (Figure 5.5b -bottom). It is the pyramidal formulation which takes into account these large displacements since the optical flow values are calculated on a coarse-to-scale manner.

We have also proposed to change the initial image used for the GVF field calculations. This affects as well the external energy formulation for the active contour approach. In (Figure 5.6- top left), it is important to select an appropriate scale, σ, when applying a Gaussian filter to the input image, otherwise the GVF vector fields will show some discrepancies near the distal part of the artery as pointed by the arrow in the figure.

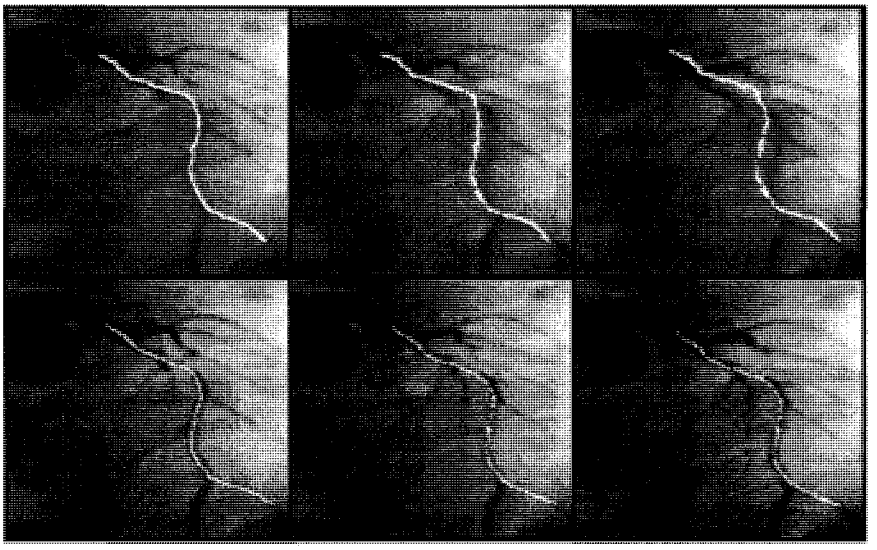


Fig 5.5a. (*Top row*) Results for traditional Lucas-Kanade Optical Flow applied to frame number 2, 3 and 4 of the first angiographic database. The technique fails when large deformations are experienced by the object being tracked. (*Bottom row*) The pyramidal implementation of Lucas-Kanade rectifies the problem for this particular angiographic sequence.

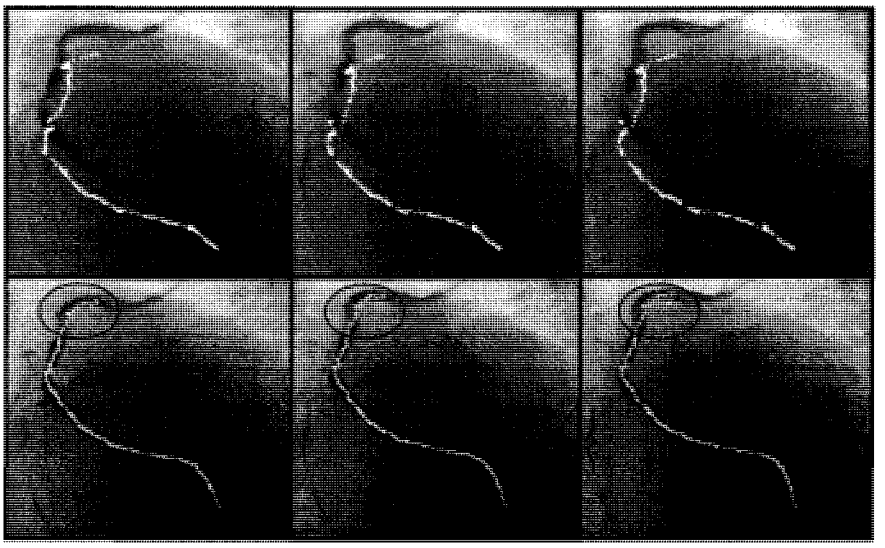


Fig 5.5b. (*Top row*) Results for traditional Lucas-Kanade Optical Flow applied to frame number 2, 3 and 4 of the first angiographic database. (*Bottom row*) The pyramidal implementation of Lucas-Kanade performs relatively well; however, it begins to falter for deformations larger than 10 pixels as shown by the elliptical region.

This might be tricky as the scale can change value depending on the amount of noise present in the dataset. Therefore, we found that by choosing our 4- step filter over all experiments (Figure 5.6- bottom left), we avoid having to determine a suitable scale for the edge map computation. The computation becomes automatic, thereby reducing simulation time. An example of the active contour convergence using the two different edge maps is shown in (Figure 5.6- right). The active contour coordinates converged properly at the distal end of the artery when using our proposed edge map versus the Gaussian convolution edge map.

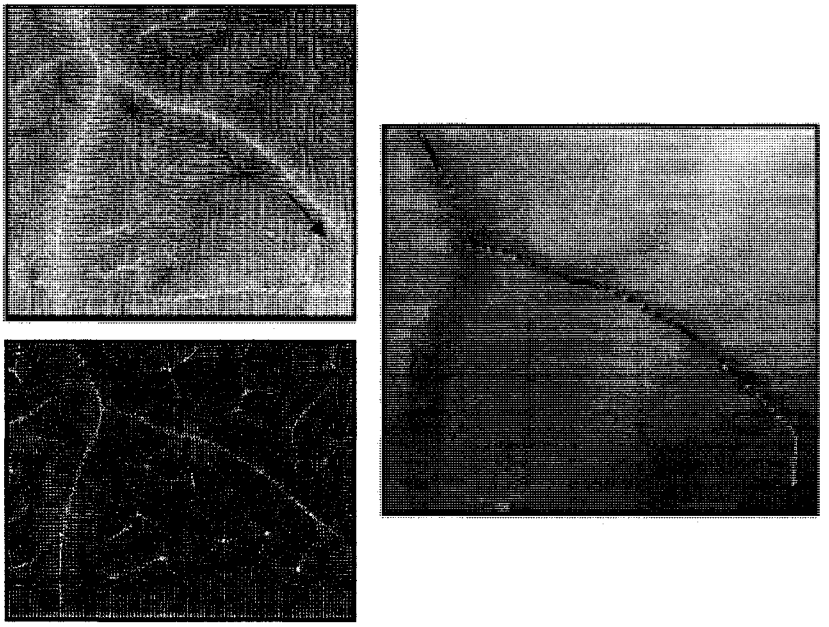


Fig 5.6. (*Left-top*) Region of interest representing distal part of coronary artery. The gradient vector flow is computed using the traditional convolution technique with a user selected scale σ. The arrow shows that the GVF fields do not converge near the distal end, but continue towards the bottom left of the image. (*Left -bottom*) By using the 4-step filter, the GVF fields converge correctly towards the distal end of the artery. (*Right*) The circle (o) centerline is obtained using the GVF field produced by the 4-step filter and depicts correctly the artery center axis, whereas the dotted line (.) centerline is a misrepresentation using the convolution image for GVF calculations.

Our algorithm correctly tracks the principal coronary arteries in 35 out of 38 images. Figure 5.7 shows typical tracking results.

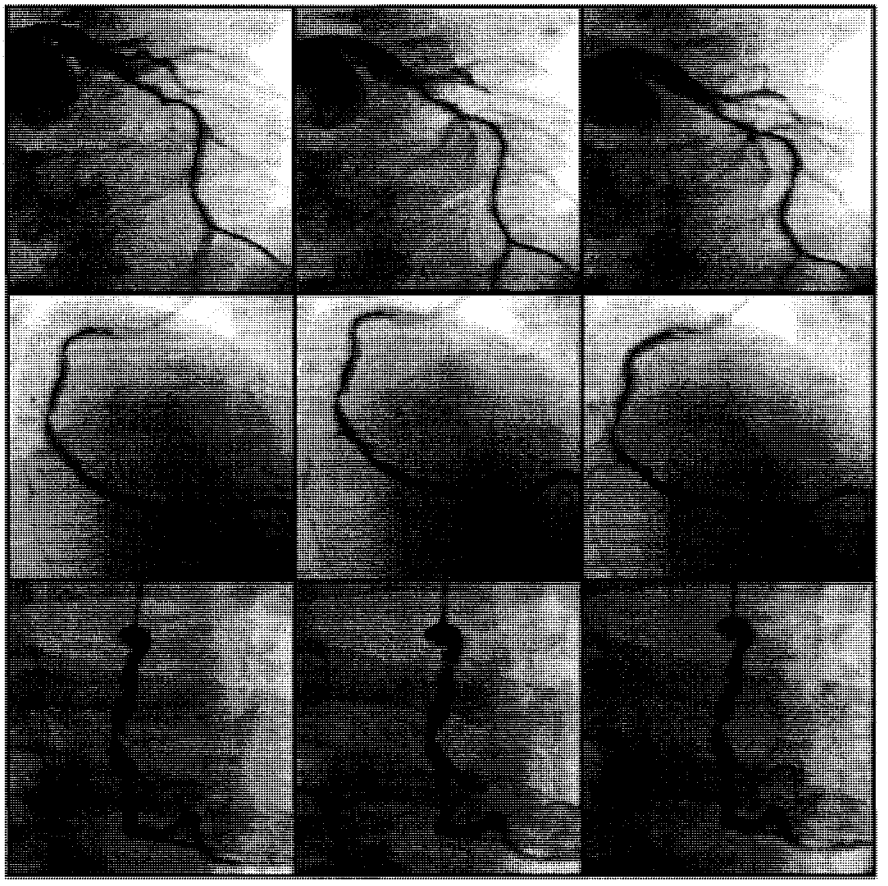


Fig 5.7. Tracking results for three X-ray angiographic databases. (Left to right columns) Respectively the 2nd, 5th and 8th image frame number.

As presented in Table 5.2, the average number of iterations for convergence using our empirically determined parameters was 12, 28 and 196 respectively for the three datasets. The last three image frames from the third dataset yielded incorrect centerline convergence as some contour coordinates were "attached" to an external structure (rib cage) and hence were not attracted to the actual coronary artery centerline. Prior to these

three images, the estimated contour coordinates were slowly being attracted to the rib cage and hence farther away from the coronary artery center axis. This is reflected in the large number of average iterations (196.1) calculated for this dataset. We noticed as well that by increasing gradually the value of the iteration step $\tau \in [0.1, 1]$, the number of iterations decreased significantly by almost half. A larger τ accelerates the convergence process, however at the cost of instability.

Table 5.2 Number of Iterations required for Interframe Convergence

Dataset #	Tracked Frames	Average number of iterations/frame
1	15	12.2 ± 3.23
2	10	27.6 ± 8.44
3	13	196.1 ± 262.57

We have proposed two improvements to traditional single view tracking algorithms. First, we have applied a multi-resolution optical flow formulation to take into consideration possible large deformations of the arteries; and second, we have modified the edge map used for the GVF calculation in order to obtain correct flow fields pointing inward towards the coronary artery center axis. When compared to the previous works, our technique is also advantageous as it only requires a two click initialization step in the first image and is completely user interaction free the rest of the procedure.

5.5 Optimal Point Matching & Curvature Constraint

To validate our proposed improvement to the biplane reconstruction process we first decided to test the method on synthetic images. Therefore, we generated two common coronary cineangiography gantry setups. The first setup represents the posterior-anterior and left lateral views (PA-LAT), defined as $\xi_{PA\text{-}LAT}$= (*SID*: 100cm, *SOD*: 50cm, *PA*:90°, *SA*:0°, u_o: 256 pixels, v_o: 256 pixels), $\xi_{PA\text{-}LAT}$= (*SID*: 100 cm, *SOD*: 50 cm, *PA*: 0°, *SA*: 0°, u_o: 256 pixels, v_o: 256 pixels). Whereas the second setup represents the left anterior and right anterior oblique views (LAO/RAO), defined as $\xi_{LAO\text{-}RAO}$= (*SID*: 100 cm, *SOD*:

50cm, *PA*: 60°, *SA*: 30°, u_o: 256 pixels, v_o: 256 pixels) and $\xi_{LAO\text{-}RAO}$= (*SID*: 100cm, *SOD*: 50cm, *PA*: 150°, *SA*: 30°, u_o: 256pixels, v_o: 256pixels). We constructed artificial 3D coronary artery points and backprojected them using the above gantry parameters. This provided a pair of biplane images. We created a total of 5 biplane image pairs for the analysis. Although Cheriet et al. proved that a minimum number of bifurcation points were needed for self calibration (a total of 30 points across 5 sets of biplane images), we had the luxury of not relying only on bifurcation points but on the artery centerline coordinates. Therefore, it is possible that in our experimentation the minimum number of 30 point matches could be obtained by using only a single pair of images. We introduced motion onto the 3D coronary points in terms of random twist angles between (-2°, +2°), as well as, random translation factors equivalent to (-0.5 cm, +0.5 cm). This motion resulted in a minimum euclidean distance of 0.488 pixels and a maximum euclidean distance of 6.19 pixels between 2D image points across consecutive images. These values were chosen so as to best represent the spatial configuration of arteries as seen in real angiographic images. Depending on the acquisition rate of the fluoroscope, the maximal interframe distance we observed between arteries in some of our coronary clinical datasets is roughly 10 pixels. Hence, we believe that the motion parameters used for the experimentation represented well the general interframe displacement of the arteries. Lastly, following the addition of this motion, we further added some normally distributed white noise $n \sim N(0,20d)$ to our 2D image points with a magnitude equal to 20 times the smallest difference between projected points, d. The above motion and noise should reflect fairly well the amount of artifacts present in real clinical data.

We now have 5 biplane image pairs which is the minimum required for proper temporal 3D reconstruction as shown by Cheriet et al. The 2D biplane points were given as input to the RANSAC method in order to obtain a first set of potential corresponding points. The outlier probability was fixed at ε=0.7 and the probability of choosing at least one sample free from outliers was set to ρ=0.99. These values are typical in literature for the RANSAC approach.

To determine whether a feature is an inlier or not, we set the inlier distance threshold, between a data point and the model used, to three sets of values: [1, 1.5, 2] pixels. Once a subset of inliers (2D correspondences) is obtained, we filter the points using the geometric curvature constraint using the curvature equations (equation 4.22) presented in the preceding chapter and solved for $N\kappa$. The optimized three dimensional curvatures, κ, were extracted accordingly for each possible corresponding pairs of matches. The above procedure is performed for each image, on a temporal basis. Lastly, we perform self calibration on the 5 biplane angiographic images using the total of the final corresponding matches. The Levenberg–Marquardt algorithm is used for optimization, iterating until the correction to the geometric parameters becomes negligible. The set of parameters and projection matrices is therefore regenerated.

Final results will focus on the number of candidates, RANSAC residual error and Sampson errors before and after the geometric constraint is applied to the potential matches. We also present the residual norm error for the self calibration procedure obtained using potential matches before and after the geometric constraint is applied. The function to be minimized is the distance between the known and optimized 2D image coordinates.

Tables 5.3 and 5.4 show the results obtained for the two synthetic simulation setups. The first column defines the biplane image pair number with the thresholds used for the inlier estimation. The following three columns present the final candidate numbers obtained, residual error and Sampson errors without having refined the final subset of correspondences obtained using RANSAC. The final three columns present results obtained after geometrical curvature constraint is applied. In both setups, a total of 1472 coordinate matches are possible. The final number of candidate points after the application of the RANSAC procedure was relatively similar for both gantry setups. We observe that the number of candidates increases as the inlier distance threshold increases from 1pixel to 2 pixels. This is natural as more corresponding points become available as the distance increases from the epipolar lines. Further, if this is true, then the number of samples in RANSAC should increase as well. However, increasing the number of

samples comes at a cost as the residual and Sampson errors also increase. The maximum residual errors and Sampson errors ranged respectively between [2.24-2.53] and [0.55-0.63] pixels for the PA/LAT views and [2.31-2.41] and [0.58-0.60] pixels for the LAO/RAO views.

The final subset of candidates was chosen for an inlier distance threshold of 1 pixel. We then applied the geometric curvature constraint to these candidates in order to refine the final amount of corresponding points. We calculated the Frenet-Serret 2D tangents and curvature on these points, and in the same fashion, the values in 3D using the gantry information. Figure 5.8 shows an example of the observed and optimized 3D curvature after solving equation (4.19) when using biplane image pair #2. We note that the circles represent the 3D Frenet-Serret curvature, whereas the points represent the optimized values from the constraint equations. Similar 3D curvature is represented by a circle and a point being within the vicinity of one another. If this is not seen, then the corresponding points lie in a cusp or zero-curvature region of the coronary artery.

Table 5.3 Point Correspondence Results: PA/LAT Simulation

	RANSAC			CURVATURE CONSTRAINT		
	# of candidates	Residual	Sampson error	# of candidates	Residual	Sampson error
Image Pair #1						
Inlier distance: 1 pixel	390	0.561	0.140	60	0.399	0.098
Inlier distance: 1.5 pixels	548	1.270	0.316	62	0.955	0.236
Inlier distance: 2 pixels	695	2.268	0.565	97	2.142	0.533
Image Pair #2						
Inlier distance: 1 pixel	422	0.609	0.151	46	0.406	0.101
Inlier distance: 1.5 pixels	572	1.370	0.341	72	1.292	0.321
Inlier distance: 2 pixels	655	2.240	0.554	112	2.015	0.500
Image Pair #3						
Inlier distance: 1 pixel	424	0.654	0.163	49	0.624	0.154
Inlier distance: 1.5 pixels	528	1.451	0.362	76	1.275	0.317
Inlier distance: 2 pixels	736	2.288	0.569	82	2.037	0.505
Image Pair #4						
Inlier distance: 1 pixel	402	0.603	0.150	51	0.444	0.111
Inlier distance: 1.5 pixels	565	1.407	0.350	81	1.009	0.249
Inlier distance: 2 pixels	697	2.338	0.583	100	1.963	0.487
Image Pair #5						
Inlier distance: 1 pixel	385	0.651	0.162	41	0.520	0.129
Inlier distance: 1.5 pixels	559	1.458	0.363	70	0.830	0.207
Inlier distance: 2 pixels	675	2.527	0.626	86	1.738	0.430

Table 5.4 Point Correspondence Results: LAO/RAO Simulation

	RANSAC			CURVATURE CONSTRAINT		
	# of candidates	Residual	Sampson error	# of candidates	Residual	Sampson error
Image Pair #1						
Inlier distance: 1 pixel	414	0.624	0.155	30	0.502	0.125
Inlier distance: 1.5 pixels	569	1.427	0.355	40	0.602	0.150
Inlier distance: 2 pixels	686	2.325	0.577	44	1.377	0.341
Image Pair #2						
Inlier distance: 1 pixel	399	0.648	0.161	39	0.555	0.138
Inlier distance: 1.5 pixels	568	1.327	0.331	47	0.997	0.247
Inlier distance: 2 pixels	675	2.351	0.582	58	1.927	0.479
Image Pair #3						
Inlier distance: 1 pixel	416	0.670	0.167	34	0.529	0.132
Inlier distance: 1.5 pixels	546	1.364	0.340	39	1.301	0.324
Inlier distance: 2 pixels	712	2.313	0.575	46	1.200	0.298
Image Pair #4						
Inlier distance: 1 pixel	383	0.644	0.160	41	0.479	0.119
Inlier distance: 1.5 pixels	556	1.424	0.355	45	0.895	0.223
Inlier distance: 2 pixels	727	2.410	0.599	57	1.839	0.454
Image Pair #5						
Inlier distance: 1 pixel	412	0.578	0.144	34	0.424	0.105
Inlier distance: 1.5 pixels	555	1.446	0.360	46	1.055	0.263
Inlier distance: 2 pixels	700	2.344	0.583	50	1.628	0.405

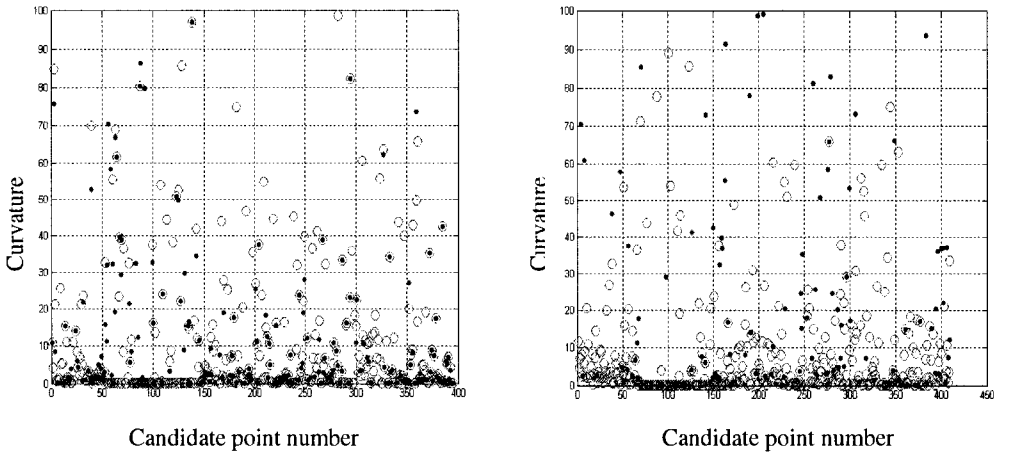


Fig 5.8. Results for the geometric curvature constraint applied to synthetic biplane image pair #2. Circles represent 3D Frenet curvature value, whereas the black points represent the optimized values for 3D curvature using the geometric curvature constraint equations. Left image shows results for the PA/LAT views and the right image provides results for the LAO/RAO setup.

The similarity measure was chosen arbitrarily to be equal to less than 0.5 units of the absolute difference between observed and optimized 3D curvatures. This threshold value was dependant on the amount of noise added to the simulations. Lastly, the value also yielded a sufficient amount of point matches required for the 3D reconstruction phase. The final number of correspondences was filtered considerably using the geometric constraint when compared to final number of RANSAC candidates. There were slightly more correspondences in the PA/LAT than the LAO/RAO simulations. One can conclude that the point of visualization of the coronary arteries will vary the number of cusps and zero-curvature points on the 2D centerlines. The maximum residual errors and Sampson errors were reduced after application of the geometric constraint and ranged respectively as [1.74-2.14] and [0.43-0.53] pixels for the PA/LAT views and [1.74-1.93] and [0.30-0.48] pixels for the LAO/RAO views.

The self-calibration of the gantry settings was then performed using the final correspondences. We introduced considerable errors in magnitude of ±6°, ±10cm and

±10 pixels to the initial values of the gantry angles, distances and image principal points. These substantial errors in approximation should represent well, and perhaps more, the mechanical and electronic errors in X-ray fluoroscopy systems found in the clinical context. Using the erroneous gantry parameters we calculated the new projection matrices for the setups and estimated the 3D coordinates using the final 2D correspondences. These approximation values were used to initiate the self-calibration process and the Levenberg-Marquardt optimization scheme was used to estimate the optimal gantry parameters. The threshold for convergence was set to 0.001. Results for the residuals are presented in Table 5.5. In both gantry setups, the residuals are smaller when considering the coronary correspondences that have been refined using the geometric curvature constraint. These results are clinically acceptable (~2mm) if we were to present reliable 3D reconstructions of the coronary artery centerlines and therefore we believe that the added curvature constraint is a good add-on tool for the RANSAC algorithm in the context of point matching.

Table 5.5 Results for self-calibration using final correspondences (in cm)

	LAO/RAO Setup	PA/LAT Setup
Norm of residual	0.262	0.159
Norm of residual using constraint	0.218	0.098
Residual	0.512	0.399
Residual using constraint	0.467	0.314

A last analysis was performed by introducing Gaussian noise to the 2D image pixels in order to calculate the 3D errors of the coronary arteries. In Figure 5.9a, we observe that as the amount of noise increases, the 3D root mean square errors (RMS) increase as well. Further, by introducing the curvature constraint for the selection of corresponding points, the RMS error decreases when compared with the RANSAC method for both gantry setups. For the LAO/RAO setup, the 3D error on average is approximately 2.8mm, whereas the average error is approximately 1.1mm for the PA/LAT setup. We

also simulated combinations of positive and negative errors in the 2D coordinates of the image coordinates for both setups. Figure 5.9b shows that errors of up to 5 pixel standard deviations do not affect noticeably the accuracy of the reconstruction process. The overall results can be explained by noting that an error of $\sigma = 3$ pixels in the image plane corresponds to approximately 2.8 mm and 1.1mm in 3D for the LAO/RAO setup and PA/LAT setup respectively. These 3D errors were generated by making use of the clinical datasets which will be presented in the following section. Further, for this same error of $\sigma = 3$ pixels, the curvature constraint method reduced the 2D retroprojection errors by a difference of $\Delta = 0.308$ pixels and $\Delta = 0.031$ pixels respectively when compared to the RANSAC method, using the LAO/RAO setup and PA/LAT respectively. From Figure 5.9a, these 2D pixel improvements correspond to an approximate 3D gain of 0.27 mm and 0.86 mm for the two views. Similarly, if the noise increased to $\sigma = 5$ pixels, then our method reduced the 2D retroprojection errors by a difference of $\Delta = 1$ pixel and $\Delta = 0.05$ pixels respectively when compared to the RANSAC method. These correspond to improvements in the 3D error of approximately 0.3mm and 0.83 mm respectively for the two gantry setting views.

Two datasets of coronary angiograms were selected for clinical validation. Image sizes were 512x512 taken from a Phillips Integris Allura Fluoroscope. The first dataset had the following gantry settings: (PA: -44.5°, SA: -1.3°, SID: 94.7 cm, u_o: 256 pixels, v_o: 256 pixels) and (PA: 45.1°, SA: -1.3°, SID: 94.7 cm, u_o: 256 pixels, v_o: 256 pixels), whereas the second dataset had the following settings: (PA: -43.8°, SA: -1.3°, SID: 93.9 cm, u_o: 256 pixels, v_o: 256 pixels) and (PA: 43.8°, SA: -1.3°, SID: 96.5 cm, u_o: 256 pixels, v_o: 256 pixels).

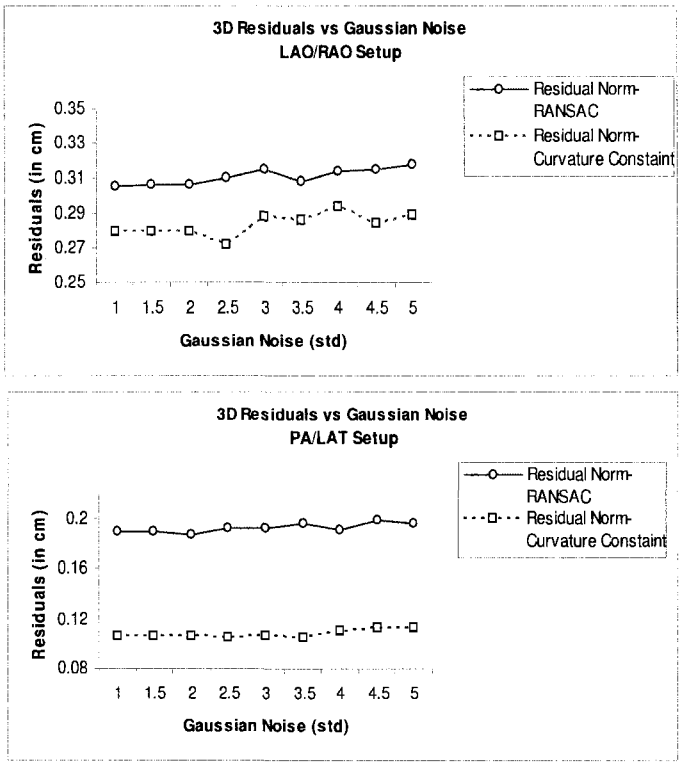
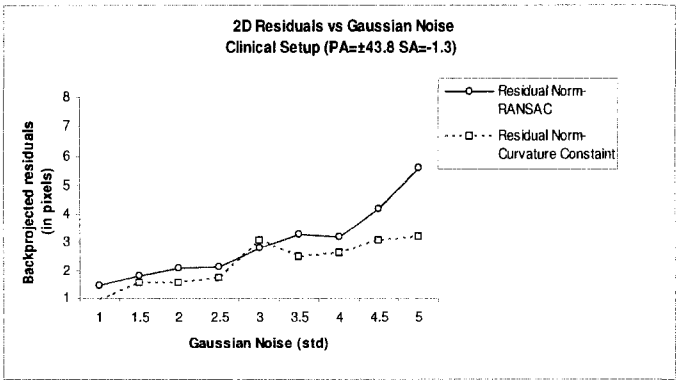


Fig 5.9a. 3D Root mean square errors for synthetic simulations. (*Top*) Gantry setting representing the LAO/RAO view and (*bottom*) the PA/LAT view. Results show that the curvature constraint yields lower 3D errors in both setups.

Principal coronary artery centerlines were first extracted for 5 biplane image pairs for each of the datasets. Our method was then implemented to globally self calibrate the system. 3D reconstruction errors using our method on the datasets are presented in Table 5.6. Mean results show amelioration in residual errors when using the curvature constraint criterion for the point correspondence refinement. After bundle adjustment the errors were 1.44 mm and 1.86 mm respectively for the two datasets.



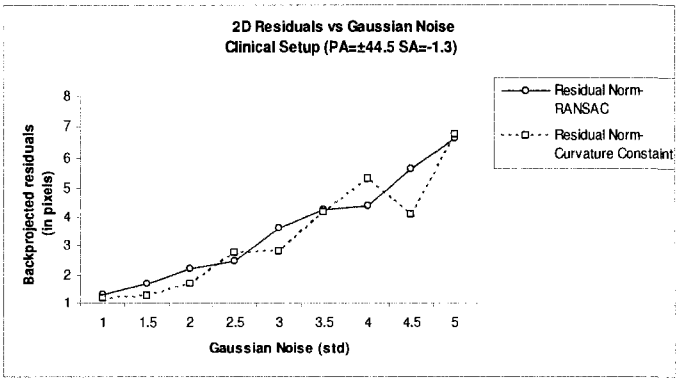


Fig 5.9b. 2D Back projection errors for clinical datasets. (*Top*) Gantry setting representing the LAO/RAO view and (*bottom*) the PA/LAT view. Results show that the curvature constraint yields lower 3D errors in both setups.

In Figure 5.10 and Figure 5.11, we provide an example on how the point correspondences were refined using our method. Using an inlier distance of 1 pixel for the RANSAC analysis, 268 candidate matches were obtained from a potential of 616. These were refined to 49 matches using the curvature constraint as seen in (Figure 5.10 - bottom). Similarly, 212 candidate matches were obtained from a potential of 753 in the image pair example of (Figure 5.11 -top). These were refined down to 36 in number after the constraint was applied. Results are promising for global self calibration of the gantry setting parameters. Temporal tracking of coronary landmarks is not a requirement in this case as point matching between individual image pairs is sufficient. The mathematical implementation is elegant and low in complexity and we believe that the method has great potential for future development.

Table 5.6 Results (average) for 2D self-calibration using clinical data

	Dataset 1	Dataset 2
Norm of residual	0.026	0.055
Norm of residual using constraint	0.021	0.034
Residual (cm)	0.161	0.235
Residual using constraint (cm)	0.144	0.186

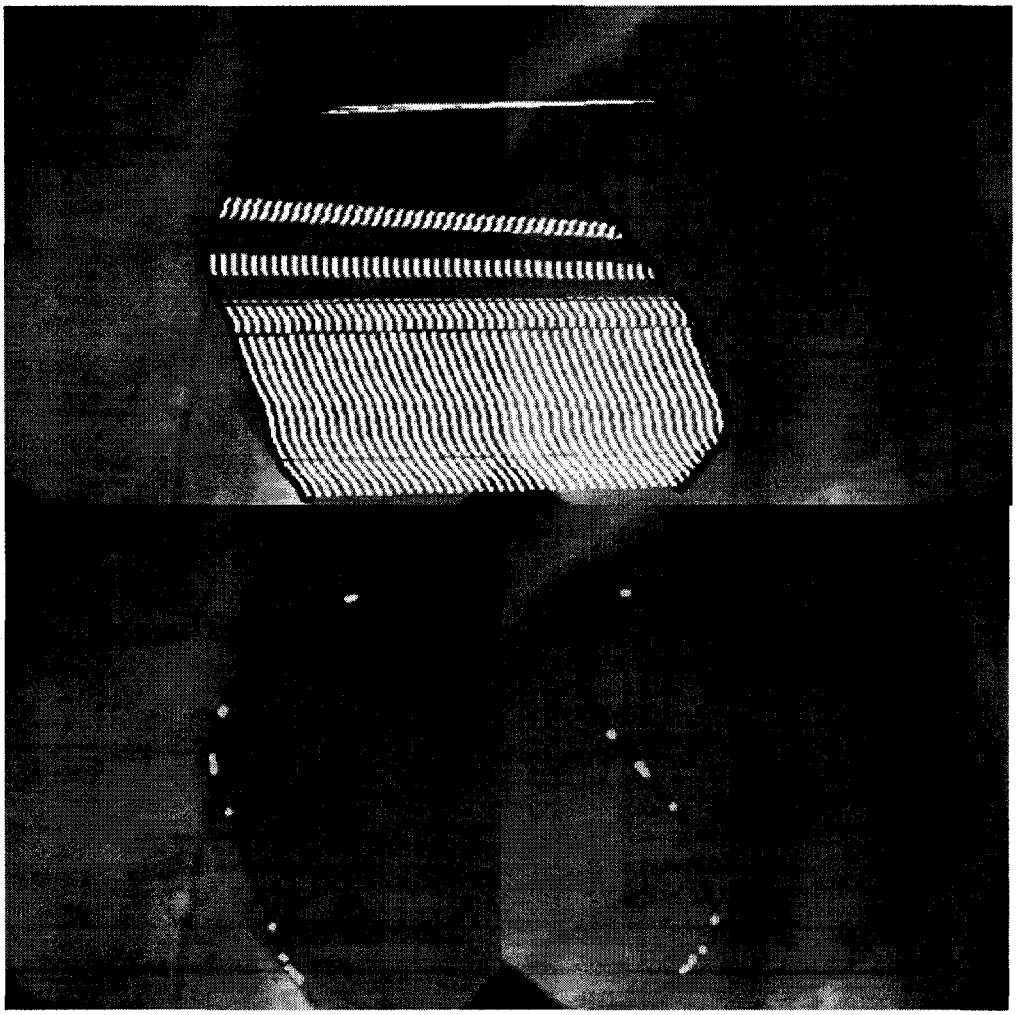


Fig 5.10. Example results for clinical dataset #1. (*Top*) RANSAC method applied on an image pair. 268 out of 616 candidates were matched. (*Bottom*) Outliers removed and curvature constraint applied to reduce the potential matches.

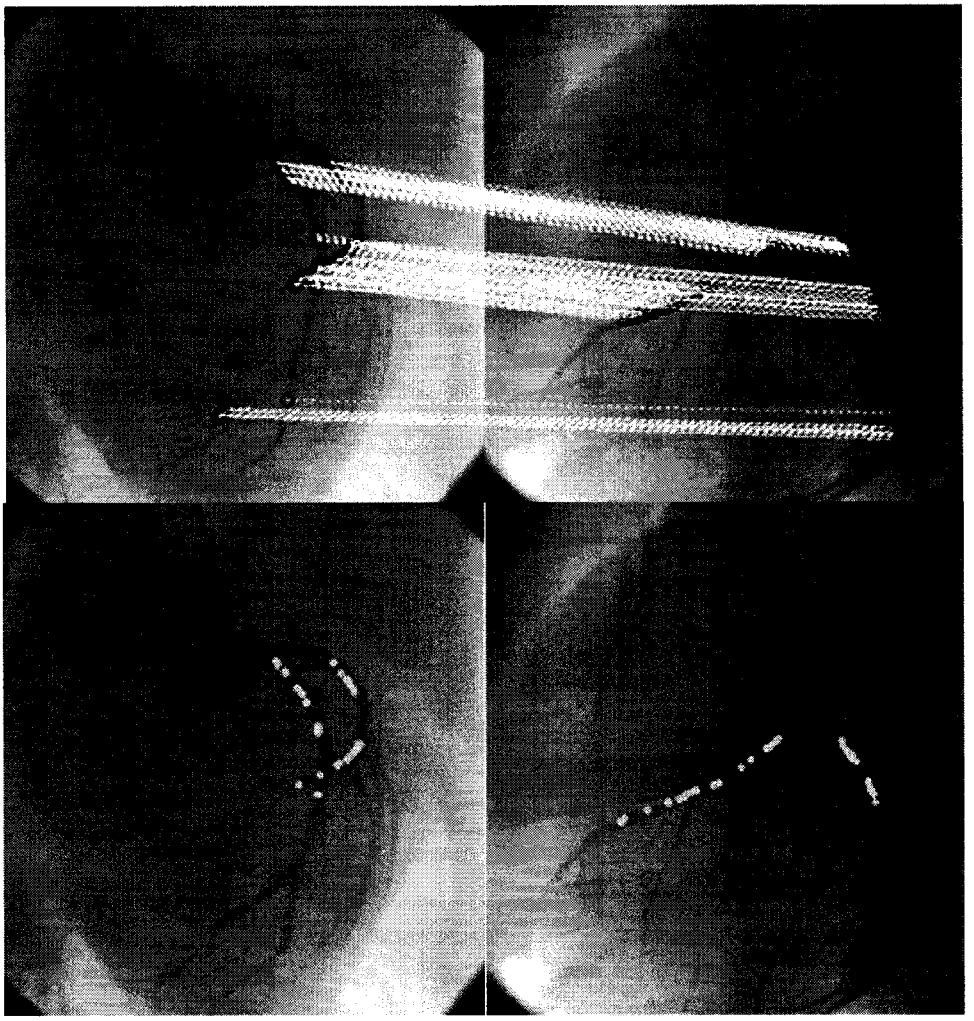


Fig 5.11. Example results for clinical dataset #2. (Top) RANSAC method applied on an image pair. (Bottom) Outliers removed and curvature constraint applied to reduce the potential matches in number.

To our knowledge, we are the first to incorporate the above tasks in the reconstruction framework of coronary artery vessels. We have segmented the coronary artery centerlines and used the centerline coordinates between biplane image pairs as potential matches. These matches were given to RANSAC and the outliers were filtered by applying an inlier distance threshold of 1 pixel with respect to epipolar lines. Results are

deemed promising and the mathematical concepts are not difficult to implement. In a clinical setting, the importance of successful 3D reconstruction is crucial for the diagnosis of coronary pathologies such as atherosclerosis, and we believe that by introducing the geometric curvature constraint the final corresponding matches will be accurate enough to provide correct 3D representations of the coronary artery centerlines.

5.6 Monoplane 3D Reconstruction

5.6.1 Helix Simulations: Rigid Motion

In order to test the feasibility of our proposed single plane reconstruction method, we performed 20 different simulations on a helix that undergoes random motion across the minimum of three consecutive images and up to a maximum of six images. We chose a helix configuration as it can reflect accurately the shape of a coronary artery centerline. The random motion was chosen to incorporate possible rotation along the three axes (bending, twist or flexion) of the 3D curve and also possible translation. Recalling that rigid motion of a 3D point can be written as:

$$\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} = [R_Z \times R_Y \times R_X] \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \tag{5.2}$$

The angles used for the simulations $\theta_x=\theta_y=\theta_z \in [0.5^o, 1^o, 1.5^o, 2^o]$. The translations used for the simulations were $T_x = T_y = T_z \in [0.5, 0.75, 1, 1.25, 1.5]$ millimeters. Thus we have 20 simulations for each possible scenario. Once again, these parameter values were chosen to reflect as best as possible the amount of interframe displacement observed from arteries in typical coronary angiograms (i.e. worse case scenario of 10 interframe pixel displacements).

The gantry setup used for this simulation will be a typical left lateral-posterior/anterior view having the same parameter values as the ones proposed in the

preceding section for the biplane reconstruction evaluation. The reason for this choice of gantry setting will become evident when we perform clinical validation. The hypothesis we make here is that we have available to us the true 3D coordinates of the helix points at a first time instant. The unknowns we need to optimize are the n-3D displacements in subsequent frames $[dx_n, dy_n, dz_n]$. Figure 5.12 shows an example of the 2D projections of the helix using the specified gantry settings. The displacement between each image frame for the left lateral view was on average 6.5 pixels, whereas the displacement between each image frame for the posterior/anterior view was 5 and 7 pixels respectively. Using orthographic projections for the approximations of dx and dy to solve our monoplane equations we obtain approximations in the range of [1.7-2.4] millimeters using the intensifier size and image size values.

As for the depth approximation dz, we can use a weak perspective projection assuming that the average depth at the first time instant will be approximately the average depth for the subsequent time instants. Thus, the Z_{avg} for the weak perspective camera model will come into play and the dz will be zero in this case. Equivalently, this can be interpreted as having the depth coordinates of the helix points in subsequent frames to be equal to the depth at the first time instant for each point. This is false as each point moves temporally, however, we rely on the theory that the average depth is a close enough approximation to initialize the optimization of our monoplane equations in hopes of converging correctly to the true 3D coordinates at subsequent time instants. We have also simulated cases in which dz had a non zero value and hence the approximations used for the depth were in the range of [0-2] millimeters.

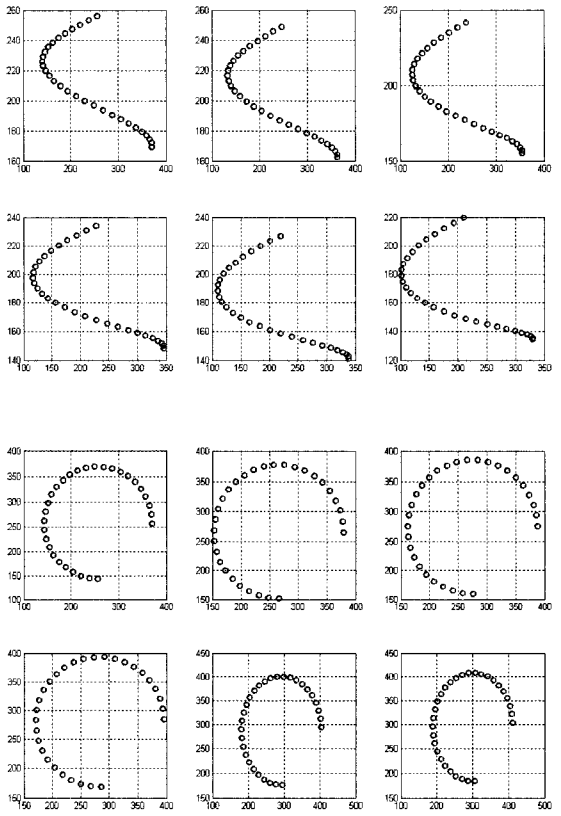


Fig 5.12. Two dimensional coordinates across six images using a 3D helix. In this case, three dimensional rotations across the three axis was set to 0.5 degrees, whereas 3D translations were set to 1.5mm in the X, Y and Z directions.

(*Top*) Left Lateral view projections across six temporal images. The average interframe pixel distance was 6.5 pixels in both the x and y directions. (*Bottom*) Posterior/Anterior view projections. Interframe pixel displacements were on average 5 and 7 pixels in the x and y directions respectively.

Table 5.7 presents results for the case where the motion helix underwent motion having a magnitude of 2° for the spatial angles and with a translation varying in distance. Our first observation is that the 3D root mean square results for our reconstruction were inaccurate if the depth approximation was equal to 2mm for both gantry setups. By assuming the average depth hypothesis results become interesting as the mean 3D reconstruction errors ranged between [1.8- 3.8] millimeters for the left lateral view and [1.3-4.0] millimeters for the posterior/anterior view. These mean values were the average results when tracking several points across the images. In total, we tracked 30 landmarks of the helix on a temporal basis.

We also present results for the minimum, mean and maximum errors of 3D depth estimation for the overall number of points tracked temporally. Let us take for example the left lateral view setup and a motion configuration of [2°, 1mm]. If we look at the table section that has as depth an initial estimation dz = 1mm, we notice that the average minimum error over the 30 landmarks tracked was 1.269 mm, whereas the maximum depth error estimated was 6.093 mm.

Before presenting results for the reconstruction process using more than three consecutive images, we will evaluate our method by using clinical data from the contraction of a left ventricle.

Table 5.7 Average results for 5 simulations using helix model (Angle: 2 degrees, Translation: 0.5-1.5 millimeters, Number of Views: 3)

Left Lateral View

		2mm depth initial error				1mm depth initial error				0mm depth initial error			
Angle (°)	Translation (mm)	3D RMS	Min Error	Mean Error	Max Error	3D RMS	Min Error	Mean Error	Max Error	3D RMS	Min Error	Mean Error	Max Error
2	0.5	6.341	2.360	5.844	9.784	5.055	1.750	4.611	7.840	3.806	1.147	3.451	6.040
2	0.75	5.804	2.177	5.274	8.91	4.446	1.509	4.054	6.971	3.251	0.938	2.919	5.412
2	1	5.162	1.885	4.699	8.027	3.835	1.269	3.494	6.093	2.716	0.765	2.393	4.838
2	1.25	4.524	1.616	4.115	7.141	3.238	1.049	2.935	5.211	2.223	0.509	1.89	4.374
2	1.5	3.888	1.373	3.535	6.252	2.66	0.8	2.382	4.483	1.8	0.325	1.458	3.856

Posterior/Anterior View

		2mm depth initial error				1mm depth initial error				0mm depth initial error			
Angle (°)	Translation (mm)	3D RMS	Min Error	Mean Error	Max Error	3D RMS	Min Error	Mean Error	Max Error	3D RMS	Min Error	Mean Error	Max Error
2	0.5	7.069	3.618	6.647	9.473	5.536	2.640	5.178	7.524	4.016	1.764	3.727	5.577
2	0.75	6.361	3.201	5.976	8.564	4.834	2.226	4.51	6.619	3.326	1.418	3.085	4.674
2	1	5.628	2.751	5.275	7.638	4.111	1.808	3.816	5.698	2.624	1.073	2.434	3.757
2	1.25	4.913	2.326	4.592	6.721	3.406	1.465	3.155	4.785	1.952	0.754	1.812	2.851
2	1.5	4.184	1.891	3.895	5.786	2.695	1.1	2.495	3.857	1.303	0.409	1.195	2.121

5.6.2 Left Ventricle Simulations: Non Rigid Motion

The clinical data for the ventricle of a sheep was obtained from [133]. In (Figure 5.13a) we can visualize 13 sonomicrometry crystals in the mitral valve and left ventricle of a sheep. Figure 5.13b represents the 2D projections of six consecutive real time instants of the contracting ventricle using the previous gantry parameters for the helix simulations. We will perform monoplane reconstruction across the six image frames and we will compare the results for this type of non rigid motion of the ventricle to the motion simulated for the helix.

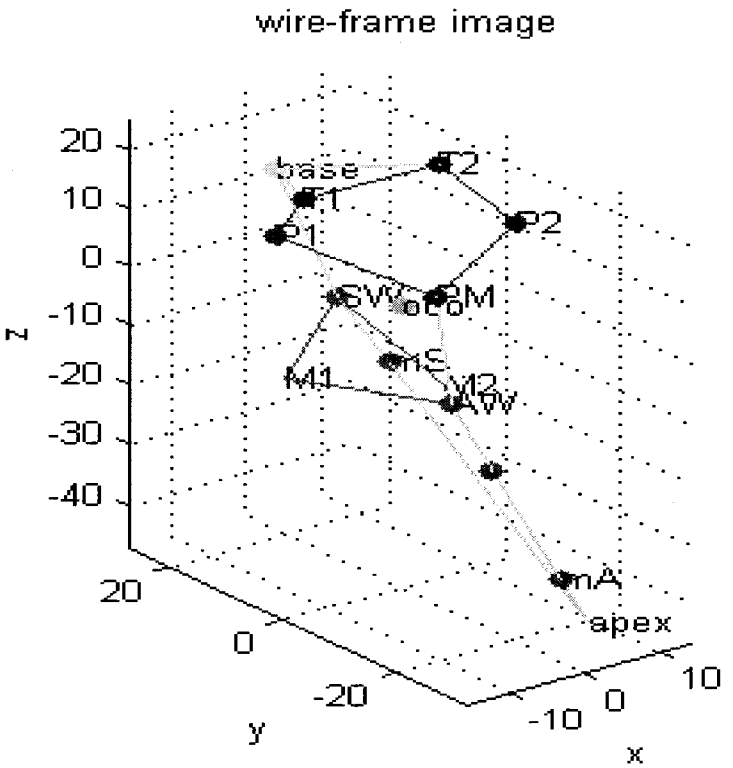


Fig 5.13a. Thirteen sonomicrometry crystals in the mitral valve and left ventricle of a sheep. Image taken from [133].

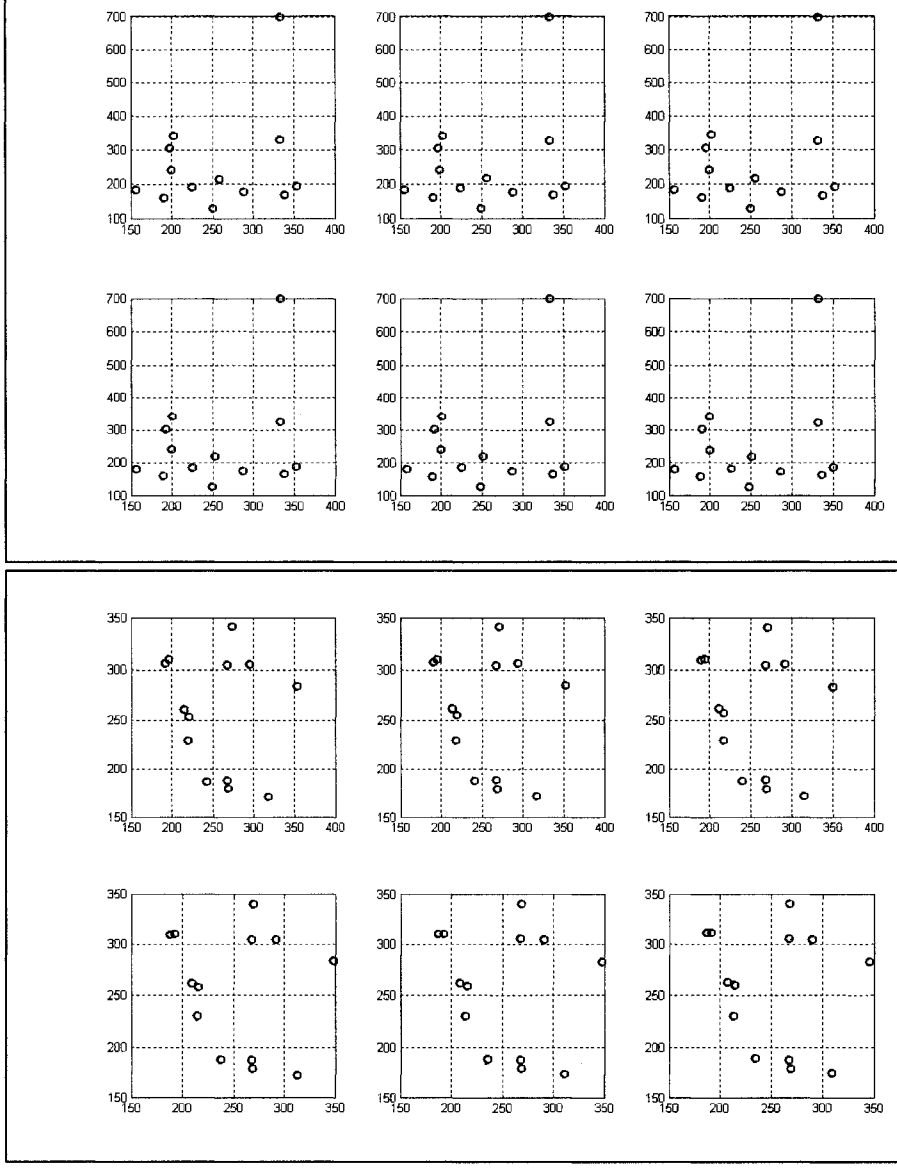


Fig 5.13b. Two dimensional projections of 13 ventricle 3D crystal coordinates across six consecutive images representing (top) anterior/posterior and (bottom) left lateral X-ray gantry configurations.

Figure 5.14a shows the best result obtained of the 20 simulations for the helix motion compared to the general simulations on the left ventricle data as shown in Figure 5.14b.

The first observation we make is that the 3D RMS error increases as we increase the number of images used for the temporal tracking of specific landmarks on either the helix curve or the sonomicrometry crystals. For an initial guess of dz = 0mm, the average 3D RMS increases from about 1 mm to 2.5 mm when tracking 29 helix landmarks across six images. However, a displacement approximation of dz = 1 mm yielded the better results when investigating rigid motion. Hence, we decided that for clinical validation we would set the depth to a 1 mm approximation for the optimization scheme. The average 3D RMS remained stable at 1 mm when tracking the 13 crystals across six images. It is to note however that since the interframe pixel distances were at most 3 pixels in the x and y directions, then the orthographic projections yielded roughly $dx = dy$ = 1 mm approximations.

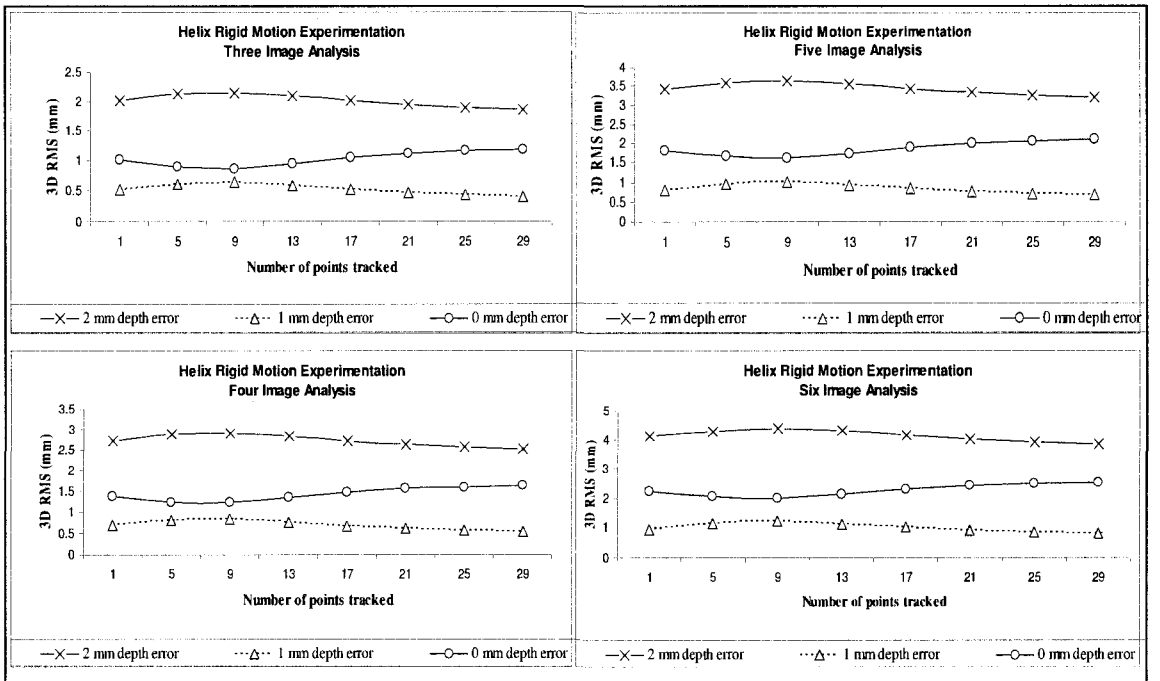


Fig 5.14a. 3D RMS Reconstruction Errors for Helix Analysis. Results for helix undergoing motion between three and six consecutive image frames.

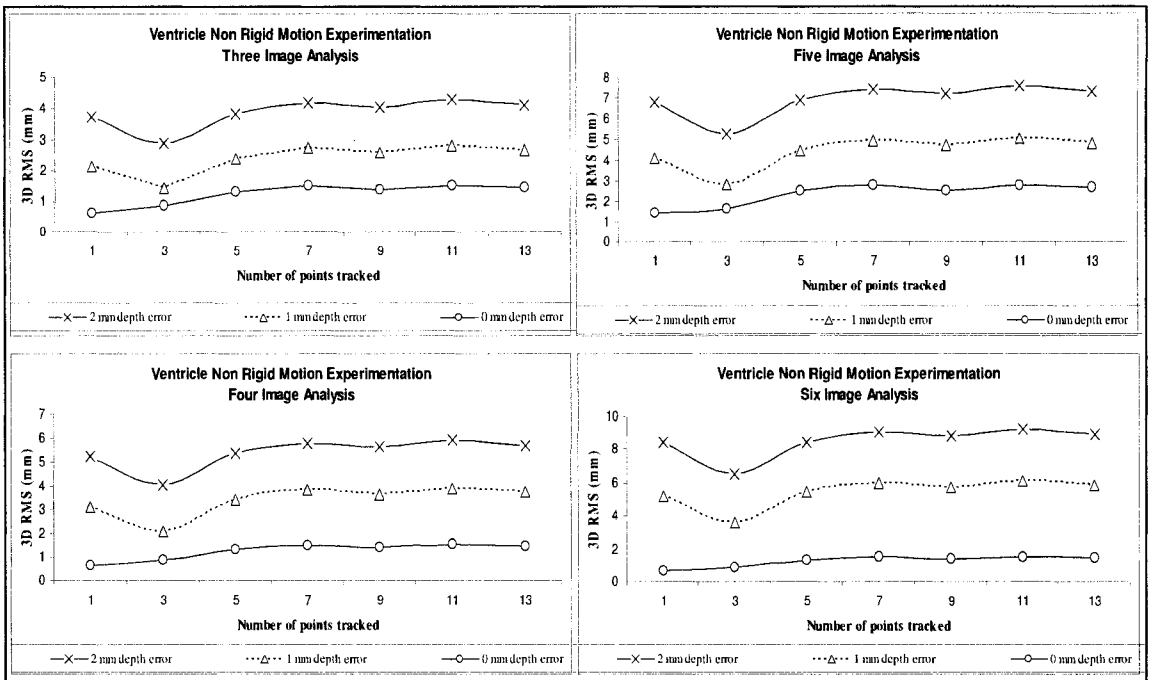


Fig 5.14b. 3D RMS Reconstruction Errors for Left Ventricle Analysis. Results for 13 markers on left ventricle undergoing non-rigid motion between three and six consecutive image frames.

This said, the clinical database from [133] confirmed that 8 consecutive images represented 1 cardiac cycle, and that the displacement was approximately 10 mm at most for one of the crystals through the eight images. Hence, this is equivalent to an average true 3D displacement of 10 mm / 8 images = 1.25 mm/image, which is close to the orthographic projection approximations we made. Nevertheless, the non rigid movement of the crystals was not recovered by the monoplane equations we proposed for the reconstruction process. This is evident when considering a depth approximation of dz= 1mm. The average 3D RMS increased from 2.5 mm to 5 mm when tracking 13 crystals across six images. We believe that twelve landmarks are sufficient in number for analysis as it presented stable reconstruction errors for the non rigid simulations. Regarding the rigid simulations, for the dz = 1mm case, twelve landmarks was close to the maximum 3D RMS error obtained.

Figure 5.15 and Figure 5.16 present results on the ability of our monoplane algorithm to recover the actual displacements $[dx_n, dy_n, dz_n]$ between consecutive images. When analyzing the minimum number of images for monoplane reconstruction, the optimization process for the helix experimentation yielded average absolute displacement errors of 0.64 mm, 0.31 mm and 1.26 mm for each of the depth approximations. As for the left ventricle analysis, the average absolute displacement errors were approximately 0.6 mm, 1.4 mm and 2.25 mm. When analyzing six consecutive images for the monoplane reconstruction process, the optimization process for the helix experimentation yielded average absolute displacement errors of 0.68, 0.29 and 1.24 mm whereas for the left ventricle analysis, the average absolute displacement errors were approximately 0.7 mm, 1.5 mm and 2.5 mm. The major conclusion is that our proposed monoplane reconstruction algorithm can indeed recover regions of motion that undergo rigid motion. Further, according to the results in Figure 5.15 and Figure 5.16, having an initial depth approximation of $dz \neq 0$ may improve recovery of the true 3D displacements.

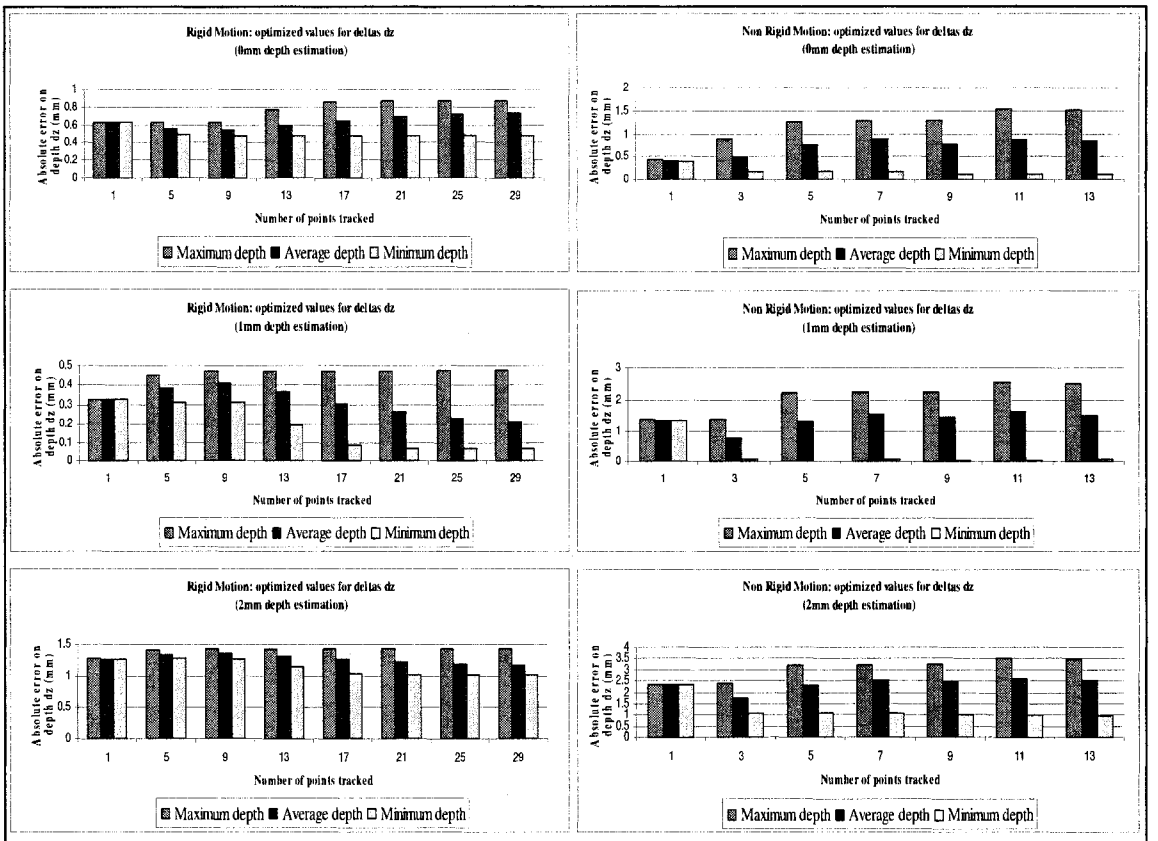


Fig 5.15. Optimized deltas, *dz*, for monoplane three view analysis in both rigid and non rigid experiments. The average absolute errors were approximately 0.64, 0.31 and 1.26 mm for the rigid motion analysis, whereas the optimized delta results were less accurate when considering global non rigid motion.

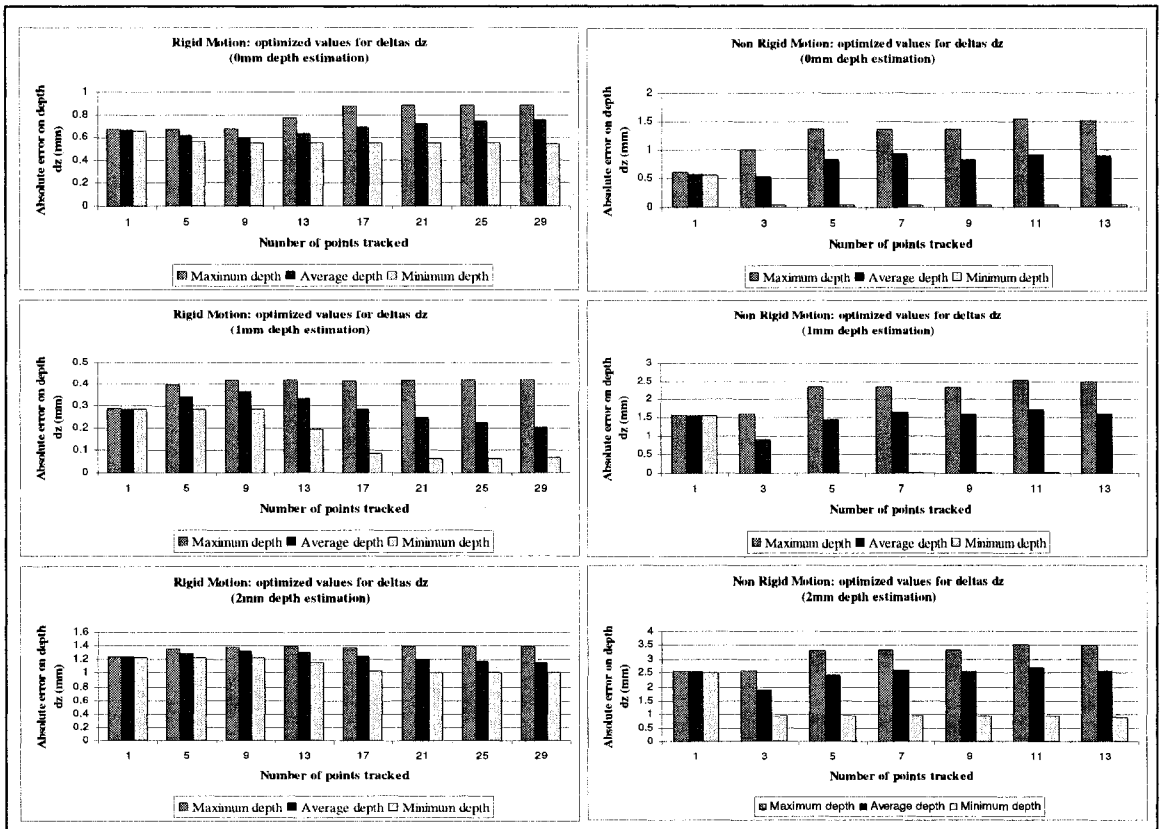


Fig 5.16. Optimized deltas, dz, for monoplane six view analysis in both rigid and non rigid experiments. The average absolute errors were approximately 0.68, 0.29 and 1.24 mm for the rigid motion analysis, whereas the optimized delta results were less accurate when considering global non rigid motion.

Figure 5.17 shows the 3D coordinates recovered by our proposed algorithm when using the left ventricle data. We observe that as the number of images increases, the reliability of the optimized results decreases significantly. This is normal as we are adding additional 3D spatial information and uncertainty to our monoplane equations. This leads us to believe that additional constraints need to be added to our equations for future analysis, if the inherent non rigid motion of the cardiac structures be recovered in its totality. Seeing how the algorithm performs generally well for rigid movements, we can target for reconstruction coronary segments such as the bifurcations, as they undergo

non-contractility type movement (rigid motion). Furthermore, we can extend our monoplane reconstruction procedure to clinical instruments such as arrhythmia ablation catheter tips as they are rigid objects.

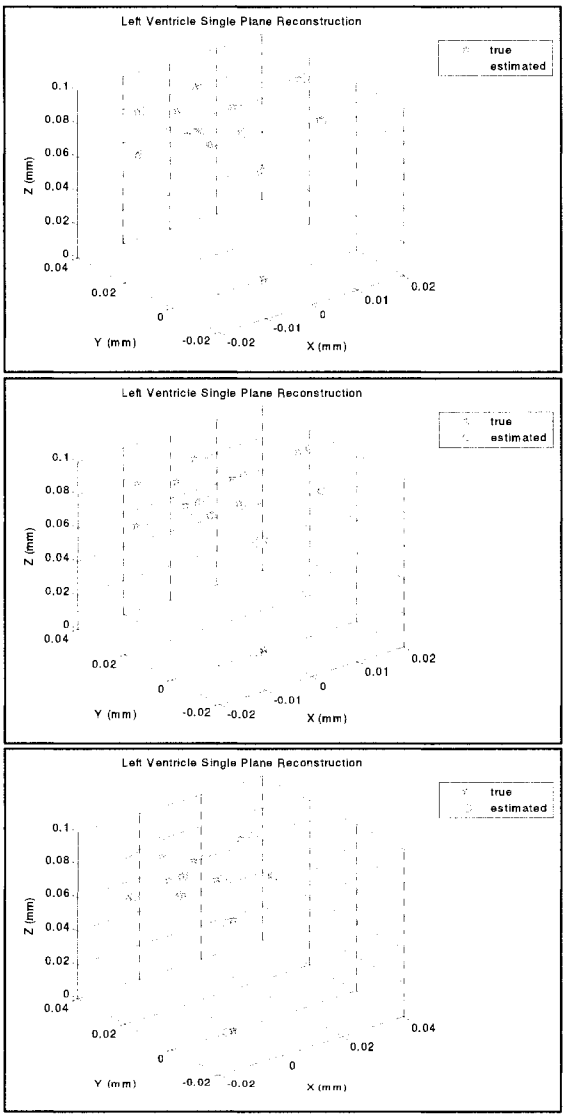


Fig 5.17. True and estimated 3D reconstruction of thirteen crystals across three images. As the ventricle deforms through the cardiac cycle, the optimized 3D positions become less accurate since we increase the amount of uncertainty regarding the true 3D configuration of the ventricle. (*Left column*) Three subsequent images following the first time instant and (*right column*) the last two image frames.

5.6.3 Clinical Validation – Ablation Catheter Rigid Motion

An animal experiment was carried out at Sacré Coeur Hospital [120]. Our experimental protocol met the institutional requirements of animal experimentation. A mongrel dog was anesthetized and laid on its right side on a fluoroscopy table (Integris Allura, Philips Inc.). A reference catheter and a pacing catheter were inserted into the right ventricle, close to the septal wall (Figure 5.18). The role of the reference catheter was to define an origin for our 3D coordinate system that is not affected by the displacement of the heart due to respiration or by the displacement of the animal or the x-ray system. The role of the pacing catheter was to produce a simple electrical activation sequence so as to validate the isochronal maps. Finally, a standard 7-French RF ablation catheter was inserted from the femoral vein into the left ventricle of the mongrel dog. During the course of the experiment, this mapping catheter was moved to 20 different sites (point-by-point) within the ventricle in order to obtain electrical and geometrical data from sufficient sites to map the activation sequence. In the end, the data collected totaled twenty posterior/anterior view biplane images, twenty left lateral view biplane images and twenty electrocardiograms.

Our monoplane reconstruction algorithm functions with the hypothesis that we have at our disposal the 3D points of cardiac objects at a first time instant $t = 0$. Therefore, we extracted the diastolic images in each gantry setting and performed biplane reconstruction using triangulation to obtain 3D points at a first time instant. Then we extracted the necessary images for the three view analysis and we also performed a five view analysis monoplane reconstruction for result comparisons. The displacement approximations in dx and dy ranged approximately between [0.5-3.5] millimeters in consequence with the interframe pixel displacements in the X-ray images. We optimized the equations using a depth estimate $dz = 1$ as initialization. We tracked twelve electrode positions from the three catheters present in the images.

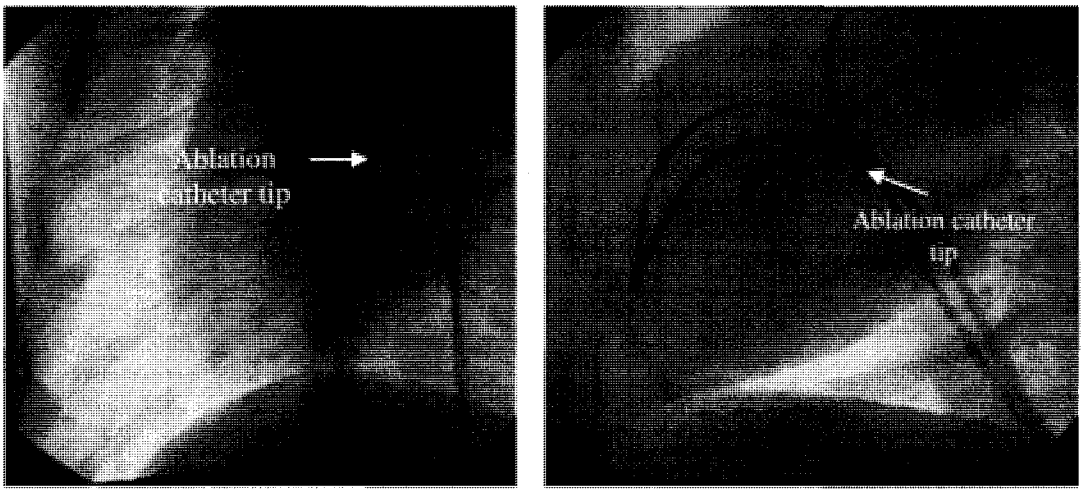


Fig 5.18. The ablation mapping catheter in the posterior/anterior view (*left*) and in the left lateral view (*right*). The catheter tip remains fixed in these diastole image frames and experience rigid motion, and hence, no twisting or bending occurs as in the non rigid case.

Table 5.8 presents the values for the 3D monoplane analysis using three consecutive images. Over the 20 datasets, we first estimated the depth for the ablation catheter tip and then calculated statistical values over the twelve electrodes used for the reconstruction process. In the posterior/anterior view, eight of the twenty datasets yielded depth errors larger than 2 mm for the catheter tip, whereas for the left lateral view six out of the twenty datasets had depth errors larger than 2 mm. Overall, the monoplane reconstruction process worked well as the average depth estimation of the catheter tip was 2.14 mm and 1.49 mm for the posterior/anterior view and left lateral views. The average maximum reconstruction error for a given electrode was 4.51 mm and 3.38 mm for both gantry setups respectively.

Table 5.9 shows the results for the five view reconstruction process. Thirteen out of twenty datasets for the posterior/anterior view produced depth estimations larger than the clinically symbolic value of 2 mm. Eleven out of the twenty datasets had depth errors larger than 2 mm for the left lateral experimentation. Overall, the monoplane reconstruction process worked well as the average depth estimation of the catheter tip was 2.68 mm and 2.34 mm for the posterior/anterior view and left lateral views. The

average maximum reconstruction error for a given electrode was 7.05 mm and 5.84 mm for both gantry setups respectively. These values reconfirm that the more images we use for the optimization process the less robust our 3D reconstruction becomes. However, it is to note that our results are about five times better than our analysis in [120] where we obtained 15 mm and 10 mm depth errors for the ablation catheter tip by using a single view. This is a substantial gain in terms of achieving a clinically valid algorithm.

Figure 5.19 and Figure 5.20 show the fusion results of the electrophysiological data and the 2D spatial positions of the ablation catheter. The left image in both figures shows the true 2D data whereas the right image shows the fusion information when estimating the 3D ablation catheter tip positions. We note that the electrophysiological data does not change significantly between the true and estimated depth analysis. For instance, Figure 5.19 shows that two electrophysiological inaccurate time values appear on the right image whereas in Figure 5.20 only one invalid time value appears.

Table 5.8 Three View Monoplane Reconstruction of Ablation Catheter

	Post/Ant *Catheter tip depth error (mm)*	**Post/Ant** *Avg. Error*	*Max Error*	**Left Lat** *Catheter tip depth error (mm)*	**Left Lat** *Avg. Error*	*Max Error*
Dataset						
1	1.163	1.896	2.934	1.218	1.493	2.351
2	2.824	3.945	5.147	0.140	0.591	1.179
3	3.283	5.718	8.403	4.125	6.090	8.230
4	1.582	2.282	3.224	0.267	0.520	1.018
5	0.957	2.338	4.304	0.466	0.342	0.942
6	1.697	2.058	2.836	3.534	5.045	6.789
7	0.277	0.517	0.950	0.326	1.718	3.386
8	4.094	6.337	8.550	0.901	1.026	1.456
9	1.420	2.052	3.264	0.768	1.344	2.089
10	0.162	1.209	2.522	0.608	0.571	0.873
11	5.365	6.806	8.573	0.287	1.429	3.045
12	1.038	0.985	1.174	2.262	3.769	5.419
13	2.830	4.342	6.410	1.349	1.849	2.688
14	1.595	2.388	3.834	1.339	2.032	3.553
15	1.401	1.980	2.758	1.416	1.675	2.558
16	4.352	5.789	7.333	4.586	6.227	8.053
17	0.981	1.197	1.689	0.110	0.354	0.769
18	2.099	3.066	4.556	2.480	3.887	5.421
19	3.936	5.763	8.074	3.563	5.312	7.364
20	1.642	2.534	3.742	0.129	0.184	0.442
Average error (mm)	2.135	3.160	4.514	1.494	2.273	3.381

Table 5.9 Five View Monoplane Reconstruction of Ablation Catheter

	Post/Ant *Catheter tip depth error (mm)*	**Post/Ant** *Avg Error*	*Max Error*	**Left Lat** *Catheter tip depth error (mm)*	**Left Lat** *Avg Error*	*Max Error*
Dataset						
1	0.333	0.958	2.218	0.416	0.790	2.472
2	0.715	2.913	8.312	3.421	7.651	11.668
3	3.802	10.641	17.025	4.681	11.015	17.860
4	2.219	5.351	9.588	1.406	3.837	6.980
5	1.498	5.379	8.841	0.824	3.804	6.703
6	2.354	7.484	12.354	3.487	7.714	11.584
7	1.005	5.717	11.205	1.147	1.615	3.166
8	4.612	11.486	18.210	3.306	7.732	12.698
9	2.025	5.433	9.367	1.111	3.797	7.798
10	1.301	5.879	10.643	1.908	4.715	8.557
11	6.580	12.283	19.630	5.803	10.953	16.797
12	3.474	11.257	18.458	3.348	9.803	15.290
13	3.328	7.013	10.140	0.373	1.492	2.683
14	2.355	6.065	10.663	2.275	5.984	10.675
15	2.481	6.096	10.425	2.572	5.988	10.360
16	6.125	11.874	19.149	3.875	9.281	16.079
17	0.921	4.012	7.986	2.212	5.265	8.640
18	3.372	7.885	12.664	0.090	2.407	6.478
19	4.474	12.175	20.975	3.832	10.476	18.314
20	0.685	1.079	1.665	0.735	2.541	5.065

Average error (mm)	2.683	7.049	2.341	5.843

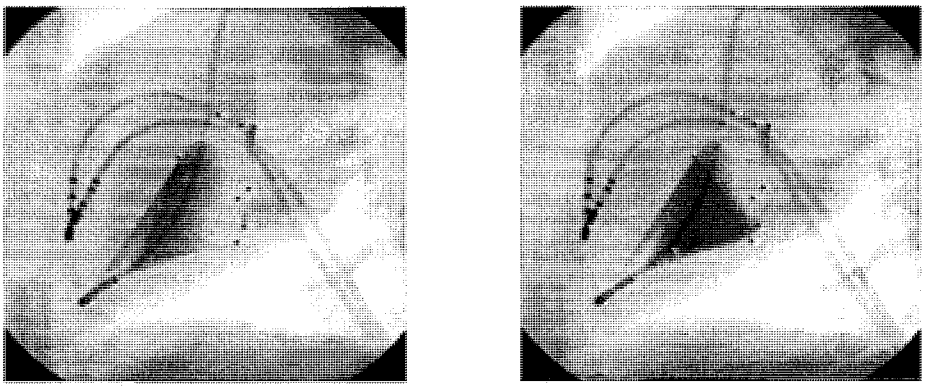


Fig 5.19. Fusion of electrophysiology on left lateral images of a mongrel dog. (Left) Visible electrophysiological data on exact depth coordinates of the ablation catheter. (Right) Erroneous electrophysiological data on estimated depth coordinates using our monoplane three view analysis.

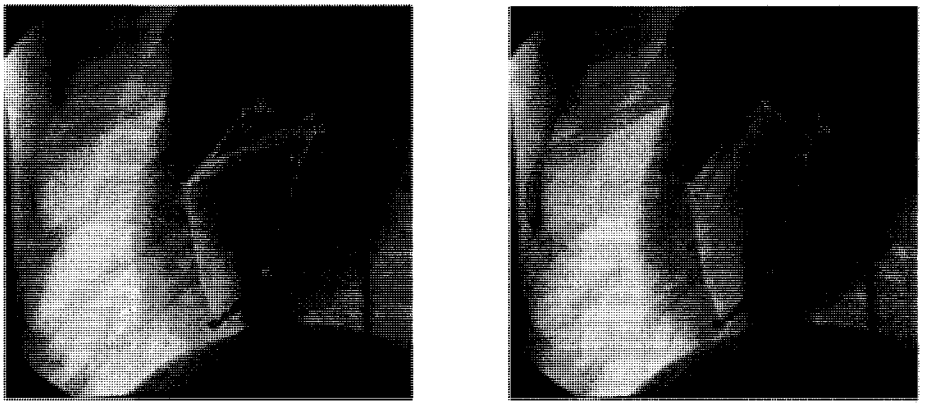


Fig 5.20. Fusion of electrophysiology on posterior/anterior images of a mongrel dog. (Left) Visible electrophysiological data on exact depth coordinates of the ablation catheter. (Right) Erroneous electrophysiological data on estimated depth coordinates using our monoplane three view analysis.

5.7 General Discussion

5.7.1 Coronary Artery Segmentation and Temporal Tracking

We developed a quick technique that selects the diastolic image given an angiographic dataset. Results confirmed that the smallest standard deviation of the difference between two X-ray images represented the diastolic phase by visually investigating the images. However, additional validation needs to be performed in order to make this method robust as we cannot simply rely on a visual inspection during clinical intervention. This is imperative as the diastolic cardiac phase would display the least motion blur and hence allow the reconstruction of a more precise coronary model.

Second, we proposed a novel 4-step filter. It would have been simple for us to use the existing Frangi et al. filter, however the challenge was to develop an automatic user friendly algorithm that would enable the interventionist to extract the diseased coronary artery without having to worry about which scales σ to select for the Gaussian convolution, in order to enhance the arteries and suppress the background by calculating the second order derivatives of the Hessian matrix. The 4-step filter enhanced the principal coronary arteries in a given angiographic image. Results showed that the filter parameters were kept constant throughout the process and all 7 principal coronary arteries were enhanced sufficiently when compared to the Lorenz and Frangi filters. However, the filter parameters have been chosen empirically and a thorough study needs to be performed to optimize them.

The resulting filtered image served as a cost image to initialize the centerline extraction process. We implemented the Fast Marching Method for its extraction. The filter parameters were fine-tuned in a way to allow the centerline extraction method to pass over possible stenosis positions (i.e. seen as a gap between the artery segment), and thus allow for a correct extraction of the centerline. We compared the cost images to those of Lorenz and Frangi and results showed that the 4-step filter performed just as well as the Frangi filter. The Lorenz filter showed a more accurate centerline in Table 5.1, however we must make mention that this result may be misrepresentative as the

artery enhanced by the Lorenz filter was much smaller in size than the Frangi and 4-step filters, since only the major eigenvalues were retained for enhancement purposes. Hence the centerline extracted using the Lorenz cost image was much closer to the actual reference centerline since the width of the artery was smaller in size.

We tracked successfully the arteries in 92% of the cases as shown in Figure 5.7. The cardiologist is required to select two extremity points of the targeted artery, in the diastolic image, and then the artery is automatically tracked in a single pass for one cardiac phase. We quantified a successful convergence if we noted the active contour at the center axis of the targeted artery. Our method incorporates a pyramidal Lucas Kanade optical flow approach that calculates the large deformations experienced by the arteries between consecutive image frames. The starting control points for the "snake" were automatically determined from the centerline coordinates. When compared to conventional tracking algorithms, our method has reduced significantly user intervention, and it does not rely on the acquisition of a 3D *a priori* model of the arteries as shown in [56]. Future work needs to focus on implementing the algorithms in C++ in order to reduce processing time as our main objective focused on providing the cardiologist with immediate assistance during interventions; hence the overall temporal tracking processing time needs to be decreased for this particular algorithm.

5.7.2 Optimal Point Correspondence

An accurate 3D model relies heavily on the successful point matching between correspondences of the structure being reconstructed. As a first analysis in this work, we believed that the centerlines well described the geometrical configuration of the arteries. The simplest manner to match the arteries in the biplane format was to locate the bifurcation points of parent branches and then perform reconstruction. Computer vision problems rely on the RANSAC method to extract relevant point correspondences in the epipolar geometry setting. We proved that these RANSAC correspondences can be further refined by taking into account the 2D/3D relationships of the landmarks. We

have added a curvature constraint that relates the 2D artery centerline curvature to its 3D Frenet-Serret curvature. The addition of such a constraint also proved to optimize the reconstruction accuracy as the potential point correspondences obtained from RANSAC were systematically filtered into an optimal subset of points. The 3D RMS errors were hence reduced when considering the biplane case. We have clinically validated our method on two angiographic clinical datasets and obtained residual result values of 1.44mm and 1.86mm for 3D reconstruction. However, it is worth noting that the RANSAC method requires a majority (~40-50%) of the data to be correct. If both false matches and severe object motion exist, RANSAC may fail, or become less attractive since many matching points due to motion are discarded as outliers. However, RANSAC yields very good results when we have a known percentage of outliers existing in the dataset. If the number of points in the dataset is unknown then the results are obtained by hit and trial or by iterating the algorithm a large number of times. Nevertheless, the final fundamental matrix estimated by RANSAC will exclude most if not all the outliers. The accuracy of the final subset of matches refined by the curvature constraint relies heavily on the accuracy of the initial pool of potential candidates acquired by RANSAC.

5.7.3 Monoplane Reconstruction

We performed both synthetic and clinical validation on our proposed single view approach. We supposed that an *a priori* 3D model of the cardiac structure or instrumentation being investigated at a first time instant was readily available. This signifies the following: (i) that a 2D/3D registration process was previously performed in order to align the 3D volume, obtained from either a CT or MRI dataset, to the corresponding anatomical structure on the 2D X-ray image or (ii) a 3D representation was first acquired using the biplane methodology and no registration be performed.

As we had no angiographic dataset at our disposal, we tested the single view approach on catheter ablation electrodes. We proved that if we track temporally a specific amount of landmarks across three angiographic X-ray images then we were able to recover most of the rigid movement of the electrodes in question. Through significant synthetic

experimentation we showed that an orthogonal or weak perspective model would be sufficient enough to initialize all 3d displacements [dx_n, dy_n, dz_n] for the optimization process. Figure 5.14 showed that under rigid motion 3D reconstruction errors were in the order of 2-3mm for our monoplane algorithm. We successfully reduced by more than 50% the overall depth estimation of the ablation catheter tip from 10mm to about 2mm using three images. However, we still encountered problems recovering the 3D displacements under non rigid motion. Figure 5.15 showed that the average absolute displacement errors were 0.64mm and 0.31mm for a dz= 0mm and dz= 1mm respectively. Compared to the computer vision field where much work has been done on single plane reconstruction, the results we obtain are encouraging. In [103], authors performed a single plane analysis on 3D rigid motion. The error obtained when calculating the optimized 3D displacements was in the range of 8%. However, the authors had more than one *a priori* 3D model unlike our case. In conclusion, we observed that by adding more image frames in the analysis, and thus more landmark spatial 2D information, we substantially decreased the accuracy for the 3D reconstruction phase.

The following question was posed: is single view reconstruction possible with no *a priori* 3D volume? With respect to our proposed monoplane equations, this signifies that the three dimensional world points (*X0, Y0, Z0*) at a first time instant are unknown. Therefore, we systematically increase the number of unknown variables from nine 3D displacements to twelve in total. Hence, other equations need be considered to solve the non linear optimization process. A first attempt was made but not outlined in Chapter 4. We exploited the two dimensional motion fields of the images calculated by the multiresolution Lucas Kanade approach. In [134], the authors derived a method for estimating 3D motion parameters of the moving object as well as the depth of the environment (Figure 5.21 and equation (5.3)). Since there are six unknowns reflecting the translational and rotational velocities on a 3D point, it is impossible to recover the 3D motion parameters given only a single image point. Thus, image velocity measurements at 5 or more image locations are the minimum requirement to solve

equation (5.3). By adding these additional equations to our proposed monoplane procedure, we thought we would arrive at a more feasible and robust method that would take into consideration the inherent non-rigid motion of the coronary arteries.

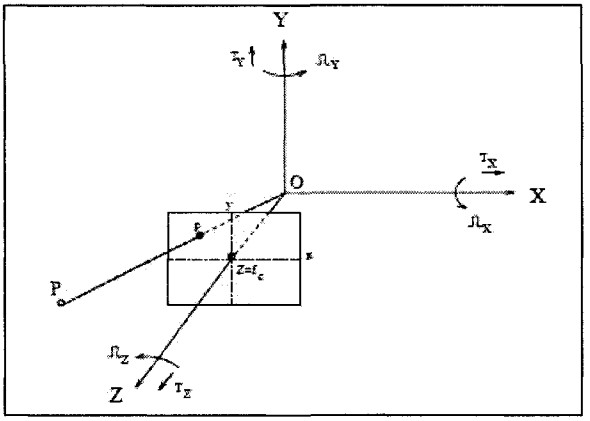


Fig 5.21. Ego Motion that states the 2D velocity of a point is related to the angular and translation motion of its 3D configurations. The coordinate system (*X, Y, Z*) is attached to the camera and the corresponding image coordinates (x, y) on the image plane are located at Z = f_c. T= (T_X, T_Y, T_Z) and Ω= (Ω_x, Ω_y, Ω_z) represent the relative translation and rotation of the camera in the scene.

$$u_{2D\ motion} = \frac{fT_x - xT_z}{Z} + f\Omega_y - y\Omega_z - \frac{xy\Omega_x}{f} + \frac{x^2}{f}\Omega_y$$
$$v_{2D\ motion} = \frac{fT_y - yT_z}{Z} - f\Omega_x + x\Omega_z + \frac{xy\Omega_y}{f} - \frac{y^2}{f}\Omega_x \qquad (5.3)$$

If the object is stationary and the camera moves, then the signs of the linear and angular velocities are reversed in equation (5.3) and, consequently, the signs of all the terms in the image motion equations are reversed as well. Preliminary results showed that equation (5.3) is a gross estimation of the movement of the coronary arteries. As mentioned, since there are initially six unknowns T and Ω, we must solve for these by using a minimum number of image pixel points. Therefore, if we define a window size surrounding a few artery pixel positions in the 2D image, then we can estimate the translation and rotational velocities within this region. The process is repeated for other window patches in order to account for all the artery pixel positions. One can conclude that an initial hypothesis was made in that within a window sized region, the artery

coordinates move with the same angular and translational velocities. This may not be the case. A first analysis was made supposing that we had an average depth $Z=Z_{avg}$ of the arteries and results were much larger than the ones described by our 3 consecutive image single plane analysis. Further hypothesis must be made to consider adding equation (5.3) to our monoplane equations.

5.8 Research Methodology Limitations

We designed methods to meet our research hypotheses and answer the specific objectives outlined in Chapter 3. However, we must outline the limitations of the proposed algorithms presented in this work. Here are some of the limitations that need to be considered when improving the proposed methodology:

1. Automatic diastolic image extraction validation was not performed using simultaneous electrocardiogram data.
2. Some of the 4-step filter parameters were determined empirically and were not fully tested to optimize their values.
3. Image segmentation was focused on the centerline geometry of the arteries instead of considering the entire vessel perimeter.
4. Monoplane 3D reconstruction was performed only on points or segments that underwent rigid motion. The non-rigid movement must be captured as well in future tests.

Chapter 6- Conclusion & Future Perspectives

This research work presented two original contributions to the scientific community. First, we proposed an optimal solution to the point correspondence problem between two non simultaneous angiographic views. Secondly, a novel single plane algorithm was introduced in order to estimate the 3D geometry of the coronary arteries, from temporal matching, across monoplane angiographic images.

The global objective of the proposed research was to elaborate and validate a method aimed at developing a potential clinical tool for the visualization and interactive manipulation of the coronary arteries, directly from the 2D X-ray angiographic images, for diagnosis and interventional purposes. This clinical tool's main purpose is to assist cardiologists in guiding them during angioplasty interventions. The goal was to develop algorithms that can be easily reproduced and applied in similar clinical procedures, such as for arrhythmia ablations, since the primary imaging modality here is the X-ray fluoroscope. Furthermore, the methodologies developed in this work are advantageous in a sense that all algorithmic parameters were selected with the intent that they remain fixed throughout the 3D reconstruction procedures. User interaction is also limited to two initial clicks by the cardiologist to extract the coronary artery centerline in the image representing the diastolic cardiac phase. Image enhancement, temporal tracking, optimal point correspondence and monoplane 3D reconstruction are achieved automatically thereafter.

We formulated two hypotheses that needed to be verified by proposing specific objectives to meet them. The first hypothesis and its specific objectives were:

Hypothesis #1: The integration of a geometrical curvature constraint improves the precision of the point correspondence and improves reconstruction accuracy.

Specific Objectives:

- Improve self calibration algorithms by introducing a novel curvature constraint that will help select valid 2D corresponding points between the non simultaneous images and make the reconstruction more precise.

- Validate constraint on biplane images first to prove the necessity of refining the point matches in a single view sequence.

Depending on the incidence angle of the fluoroscope we cannot assume that a minimum number of bifurcations would be visible in the 2D images for the point matching process. Therefore it was imperative to find other matches along the coronary artery centerline. To meet the specific objectives and validate our first hypothesis we performed both synthetic and clinical validation using centerline correspondences. Results show that for synthetic experimentation, with added gaussian noise of up to 5 pixels standard deviations, using the left-right anterior oblique (LAO/RAO) and posterior-antérior lateral (PA/LAT) viewing angles, the average 3D RMS errors are respectively 3.1 mm and 1.9 mm for the RANSAC method, and decreased to 2.8 mm and 1.1 mm by using our curvature constraint methodology. Similarly, clinical validation was performed on two angiographic datasets. By using our novel curvature constraint to refine the point correspondences between both views, the 3D RMS errors decreased from 1.61mm to 1.44mm in the first dataset and from 2.35mm to 1.86mm in the second dataset.

The second research hypothesis and its specific objectives were:

Hypothesis #2: It is possible to recuperate the 3D geometry across a single plain angiographic sequence, for clinical assistance, by using an *a priori* 3D model of the structure of interest.

Specific Objectives:

- Develop a novel automatic 2D segmentation algorithm to enhance healthy or diseased (stenosis) coronary arteries in a first image representing the diastolic phase.

- Provide the interventionist with the ability to select a specific coronary artery and investigate its motion by extracting its 2D centerline and automatically temporal tracking it across all the angiographic images in a cardiac cycle.

- Perform monoplane 3D reconstruction by implementing a novel self calibration algorithm using a single angiographic view and an *a priori* 3D model.

In order to achieve monoplane reconstruction, the first two specific objectives needed to be developed. A two-click method to extract the artery centerline was implemented and its position was temporally tracked automatically across a monoplane angiographic sequence. We proposed that the required a priori 3D model of the coronary arteries be obtained using the diastolic images from the two views obtained by rotating the C-arm fluoroscope on two occasions. The minimum number of single plane temporal images required for monoplane 3D reconstruction is three. Considering that monoplane 3D reconstruction has never been performed previously we developed equations that would lead to the estimation of the 3D geometry of the arteries across the minimum number of X-ray images. To meet the specific objectives and validate the second hypotheses we once again performed synthetic experimentation and clinical validation.

We constructed a 3D helix in space and applied rigid motion to it across six different time instants. We then reprojected those six sets of 3D points on 2D images

using projection matrices that represented the posterior-anterior view. To experiment on non-rigid moving objects, we used 3D landmarks from a sheep's ventricle and reprojected those coordinates onto six 2D images. We then optimized our monoplane equations and estimated the 3D geometry of the helix and ventricle landmarks across the six time instants. Results showed that by assuming depth errors of at most dz= 1mm, the 3D RMS errors for the helix increased from about 1mm to 2.5mm if we systematically increase the number of X-ray images from three to six. Similarly, by assuming depth errors of at most dz= 1mm, the 3D RMS errors for the ventricle landmarks increased from about 2.5mm to 5.5mm if we increase the number of X-ray images from three to six. These experiments showed that the monoplane equations developed recuperate well the 3D coordinates of objects that undergo only rigid movement. Also, by using depth approximations of dz=2mm, the optimization begins to decrease in precision since 3D RMS results are about 4mm and 8mm when using the helix and ventricle landmarks respectively. We clinically validated the monoplane approach on catheter electrode datasets to treat arrhythmia. The average 3D depth error, across the 20 datasets, when estimating the depth of the tip electrode of the ablation catheter was 2.14mm and 1.49mm when using a minimum of three X-ray images, posterior-anterior view and left lateral view respectively.

We believe that we have met our research objectives by performing sufficient experimentation to validate our proposed methodologies. As there is room for improvement in certain aspects of our research work, we follow by outlining possible solutions to remedy the various problems that still need to be addressed in our research.

In terms of future work regarding two view reconstruction there are two avenues that need to be further investigated. First, we should consider matching the artery contour coordinates instead of using only the centerline points. This is important for modeling purposes if an eventual biomechanics study were to be performed for pre and post-operative studies. Having the dynamic information of the artery walls would provide a

more complete evaluation of the progression of stenosis and follow-up of stent placement near the obstructed site. Nevertheless, performing point correspondence between vessel contours remains problematic due to the difficulties of having multiple matches for multiple perimeters (Figure 6.1).

Therefore, we can possibly consider modeling the 3D centerlines of arteries by implementing a NURBS (Non Uniform Rational Basis Spline) formulation on both the centerlines in the 3D space and in the 2D images. The advantage of considering NURBS is that they remain invariant to perspective projection. Hence, there may be a possibility of facilitating point correspondence. Furthermore, since our method only reconstructed rigid segments, we can complete the full scale 3D model by passing a NURBS through the monoplane reconstructed rigid points in order to fill the missing centerline information. Also, length preservation can now be achieved if the artery is modeled as a spline since one can easily calculate the 3D length of the reconstructed artery and correlate it to those coordinates in the 2D images to determine the 2D length of the artery.

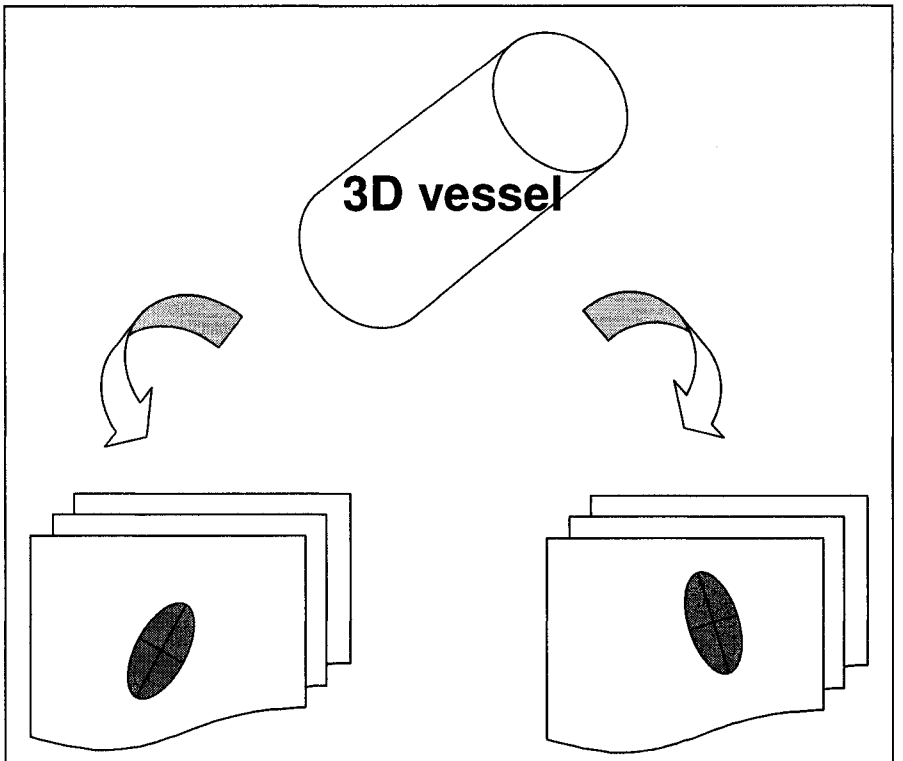


Fig 6.1. Matching 3D/2D vessel perimeters might be problematic when performing point correspondence. Therefore, a geometrical representation that remains invariant to perspective projection should be considered.

Nevertheless, the filter parameters were determined empirically and a more thorough analysis must be made to optimize the parameters for the 4 filters used. We suggest adaptive image fine tuning of the parameters. If local characteristics of the data can be extracted and correlated with the filter parameters then we can deem these optimal for algorithmic validation. Regarding the automatic diastolic image frame extraction, acquiring ECG's and synchronizing them with the images is an essential validation step to our proposed methodology. Recalling that our technique chose the diastolic image frame number that had the smallest standard deviation assigned to it, it would be imperative to verify that the actual frame number falls directly on the ECG segment representing the diastolic cardiac phase.

Furthermore, there is great potential to automatically extract the coronary artery contours by implementing a novel level set method introduced in [135]. Figure 6.2 shows the potential of such a method. Nevertheless the level set formulation relies greatly on the initial image quality and background uniformity. By extracting the contours of the arteries, we can then re-apply our tracking algorithm throughout a cardiac cycle and this will lead to a more complete reconstruction of the coronary artery geometry.

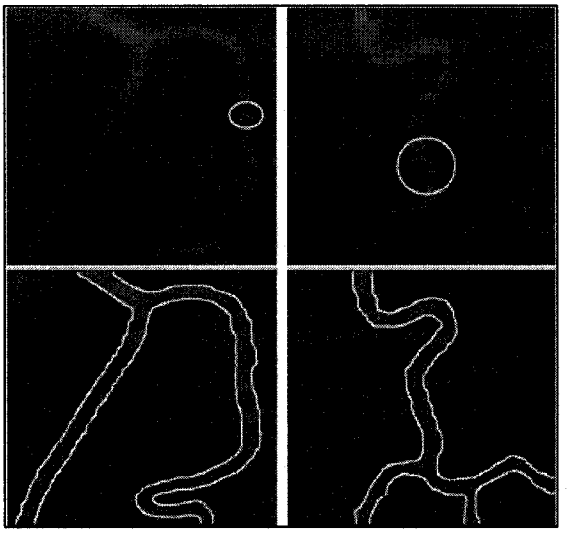


Fig 6.2. Implicit Active Contours Driven by Local Binary Fitting Energy by [135]. Level set method uses a kernel to extract local image information for the segmentation process.

Lastly, the GVF formulation failed when the active contour was far from the desired center position of the targeted artery. Although one of the advantages of the GVF technique is its large capture range, it would fail in some cases. The future improvement of our tracking algorithm should revolve around the implementation of a vector field convolution (VFC) algorithm. The VFC external force can be broken down into two independent steps: the formation of edge map from the image, and the computation of the external force from the edge map. This novel external force for active models is calculated by convolving a vector field with the edge map derived from the image. Like

the GVF technique, instead of being formulated using the standard energy minimization framework, VFC snakes are constructed by way of a force balance condition. The novel static external force has not only a larger capture range and ability to capture concavities, but also reduced computational cost, superior robustness to noise and initialization, and flexibility of changing the force field.

Regarding monoplane reconstruction, this problem is an original one to tackle in the medical field. The economic and social impact of such a breakthrough would be significant. Most hospital settings are equipped with a monoplane C-arm fluoroscope that needs to be rotated twice in order to acquire orthogonal views for 3D reconstruction. However, during clinical interventions the cardiologist more often relies only on a single view provided by a specified incidence angle of the X-ray fluoroscope. The reason for this is that the acquisition time is reduced in half by avoiding rotating the fluoroscope a second time. Also, clinical intervention time is diminished and radiation exposure is reduced. It is our obligation to provide robust clinical tools to assist the cardiologist during these interventions to facilitate navigation of catheters and guidewires and quicken diagnosis as well.

Our proposed monoplane algorithm can be improved in a number of ways. First, the addition of constraints to our equations needs to be considered. Since we modeled the coronary arteries as a spatial 2D and 3D curve during the optimization phase, then in theory each reconstructed 3D point must have its 3D tangent perpendicular to the 3D normal (Figure 6.3).

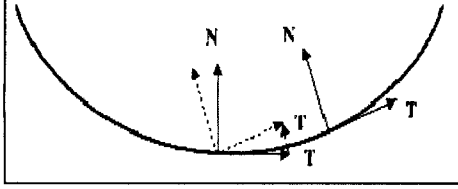


Fig 6.3. A Frenet-Serret constraint that defines orthogonality between the 3D normal and tangent.

Another additional constraint might be a length preservation criterion in which the 3D artery centerline lengths remain relatively similar to its 2D projections. This is vital when considering a registration framework between a 3D volume and its 2D image projection. If CT or MRI technology is available, then a last possible solution would be to acquire a second volume, perhaps at the systolic cardiac phase, in order to calculate approximately the amount of 3D deformation experienced by the coronary structure between diastole and systole. For example, if there are a total of 10 angiographic images representing the cardiac cycle and we determined that the 3D points moved a total of 5 millimeters when subtracting both volumes, then in consequence we can assign more precise estimations for the displacements in three dimensions (dx, dy, dz), instead of assuming orthographic or weak perspective projection camera models.

In this 21st century, there is a continuing necessity for academic researchers to work in conjunction with clinical cardiologists in order to continually improve existing technologies and establish higher standards of cardiac medical interventions (i.e. bypass surgery, angioplasty, and arrhythmia ablations to name a few) with respect to quality and time. By reducing overall intervention time with the aid of an accurate monoplane reconstruction approach, we will also reduce the stress on patients and their exposure to radiation from the fluoroscopy X-ray system. Also, a quick and efficient cardiac intervention usually means that the patient may leave the hospital sooner and this is beneficial in reducing backlog of cardiac interventions and waiting time for patients to be hospitalized. We conclude this work by stating that technological assistance to clinical interventions may be facilitated by ongoing development of robust mathematical and image processing formulations that describe motion analysis and geometrical spatial configurations.

BIBLIOGRAPHY

[1] American Heart Association. 2007 Heart and Stroke Statistical Update, American Heart Association. Available at: http://www.americanheart.org/downloadable/heart/1166711577754HS_StatsInsideText.pdf. Accessed July 1, 2008.

[2] Angioplasty.org. The most popular Angioplasty site on the web. Section: Devices: Coronary Diseases. 2004. Available at: http://www.ptca.org/devices.html. Accessed July 1, 2008.

[3] L. O'Connor, M.D. Southern California Cardiovascular Consultants. Available at: http://www.socalcardiology.com/intervention.html. Accessed July 1, 2008.

[4] A. Colombo, P. Hall, et al. Intravascular stenting without anticoagulation accomplished with intravascular ultrasound guidance, *Circulation*, 91(2), 1676-1688, 1995.

[5] Philips, CardioVascular X-Ray: Integris Allura. Available at: http://www.medical.philips.com/main/products/cardiovascular/products/allura9/. Accessed July 1, 2008.

[6] C. Lorenz et al. Multi-scale Line Segmentation with Automatic Estimation of Width, Contrast and Tangential Direction in 2D and 3D Medical Images. *CRVMed-MRCAS Proceedings*. Springer Verlag, p. 233-242. 1997.

[7] A. F. Frangi, W. J. Niessen, R. M. Hoogeveen, T. Van Walsum, and M. A. Viergever. Model-based quantitation of 3D magnetic resonance angiographic images. *IEEE Trans. Med. Imag.*, vol. 18, pp. 946–956, 1999.

[8] L. Yingchao, L. Yue, and W. Yongtian. Novel algorithm for centerline extraction of coronary arterial tree in coronary angiographic projections. *IEEE/ICME International Conference on Complex Medical Engineering*, p 830-5, 2007.

[9] S-J. Zhou, W-f Chen, Q-J. Feng,J-G. Zhang and Y-T. Wang. Extracting the coronary artery tree in angiographic projections based on probability tracking. *Acta Electronica Sinica*, v 34, n 7,1270-4, July 2006.

[10] J. Meunier and C. Bellemare. Structural method for tracking coronary arteries in coronary cineangiograms. *Proceedings - 1st Canadian Conference on Computer and Robot Vision, Proceedings - 1st Canadian Conference on Computer and Robot Vision*, p 378-384, 2004.

[11] K. Haris, S.N. Efstratiadis, N. Maglaveras, J. Gourassas, C. Pappas and G. Louridas. Automated coronary artery extraction using watersheds. *Computers in Cardiology*, p741-744, 1997.

[12] T.V Nguyen and J. Sklansky. Computing the skeleton of coronary arteries in cineangiograms. *Computers and Biomedical Research*, v 19, n 5, 428-44, Oct. 1986.

[13] T. Tozaki, Y. Kawata, N. Niki, H. Ohmatsu, and N. Moriyama. 3-D visualization of blood vessels and tumor using thin slice CT. *IEEE Nuclear Science Symposium and Medical Imaging Conference*, vol. 3, pp. 1470-1474, 1995.

[14] Y. Kawata, N. Niki, and T. Kumazaki. An approach for detecting blood vessel diseases from cone-beam ct image. *IEEE Int. Conf. on Image Processing*, pp. 500-503, 1995.

[15] N. Niki, Y. Kawata, H. Satoh, and T. Kumazaki. 3D imaging of blood vessels using x-ray rotational angiographic system. *IEEE Med. Imaging Conf.*, vol. 3, pp. 1873-1877, 1993.

[16] E. Sorantin, C. Halmai, B. Erdohelyi, K. Palagyi, L. Nyul, K. Olle, B. Geiger, F. Lindbichler, G. Friedrich, and K. Kiesler. Spiral-ct-based assessment of tracheal stenosis using 3-d-skeletonization, *IEEE Trans. on Med. Img.*, vol. 21, pp. 263-273, March 2002.

[17] S. Aylward, S. Pizer, E. Bullitt, and D. Eberl. Intensity ridge and widths for tabular object segmentation and registration. *Wksp on Math. Methods in Biomed. Image Analysis*, pp. 131-138, 1996.

[18] E. Bullitt and S.R. Aylward. Analysis of time-varying images using 3d vascular models. *Proc. Applied Imagery Pat. Recog. Works.*, pp. 9-14, October 2001.

[19] D. Guo and P. Richardson. Automatic vessel extraction from angiogram images. *IEEE Computers in Cardiology*, vol. 25, pp. 441-444, 1998.

[20] H. Schmitt, M. Grass, V. Rasche, O. Schramm, S. Haehnel, and K. Sartor. An x-ray-based method for the determination of the contrast agent propagation in 3-d vessel structures. *IEEE Trans. on Med. Img.*, vol. 21, pp. 251-262, March 2002.

[21] J. F. OBrien and N. F. Ezquerra. Automated segmentation of coronary vessels in angiographic image sequences utilizing temporal, spatial structural constraints. *Proc. SPIE Conf. Visualization in Biomed. Computing*, 1994.

[22] P.J. Yim, P.L. Choyke, and R.M. Summers. Gray-scale skeletonization of small vessels in magnetic resonance angiography. *IEEE Trans. on Med. Img.*, vol. 19, pp. 568-576, June 2000.

[23] S. Eiho and Y. Qian. Detection of coronary artery tree using morphological operator. *IEEE Computers in Cardiology*, vol. 24, pp. 525-528, 1997.

[24] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *NM*, vol. 66, pp. 1-32, 1993.

[25] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: A level set approach. *PAMI*, vol. 17, pp. 158-175, February 1995.

[26] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulation. *JCP*, vol. 79, pp. 12-49, 1988.

[27] J. A. Sethian. A fast marching level set method for monotonically advancing fronts. *NAS*, vol. 93, pp. 1591-1595, 1996.

[28] O. Wink et al. 3D MRA Coronary Axis Determination using a Minimum Cost Approach. *Magnetic Resonance in Medicine*. 47(6):1169:1175, 2002.

[29] S.D. Olabarriaga, M. Breeuwer, W.J. Niessen. Evaluation of Hessian-based filters to enhance the axis of coronary arteries in CT images. *CARS*, 1191-1196, 2003.

[30] Q. Lin. PhD Dissertation. Enhancement, Extraction, and Visualization of 3D Volume Data. Institute of Technology. Linkopings University, Sweden. 2003.

[31] M. Kass, A. Witkin, and D. Terzoopoulos. Snakes: Active contour models. *Int. J. of Comp. Vision*, vol. 1, pp. 321-331, 1988.

[32] C. Molina, G. Prause, P. Radeva, and M. Sonka. 3D catheter path reconstruction from biplane angiograms. *SPIE*, vol. 3338, pp. 504-512, 1998.

[33] A.K. Klein, F. Lee, and A.A. Amin. Quantitative coronary angiography with deformable spline models. *IEEE Trans. on Med. Img.*, vol. 16, pp. 468-482, October 1997.

[34] A. Yilmaz et al. Object Tracking: A Survey. *ACM Computing Surveys*, Vol. 38, No. 4, Article 13, December 2006.

[35] T. Broida and R. Chellappa. Estimation of object motion parameters from noisy images. *IEEE Trans. Patt. Analy. Mach. Intell.* 8, 1, 90–99, 1986.

[36] J. Shi and C. Tomasi. Good features to track. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 593–600, 1994.

[37] D. Huttenlocher, J. Noh and W. Rucklidge. Tracking non-rigid objects in complex scenes. *IEEE International Conference on Computer Vision (ICCV)*. 93–101, 1993.

[38] K. Sato and J. Aggarwal. Temporal spatio-velocity transform and its application to tracking and interaction. *Comput. Vision Image Understand.* 96, 2, 100–128, 2004.

[39] J. Meunier, M. Bourassa, M. Bertrand, M. Verreault, and G. Mailloux. Regional epicardial dynamics computed from coronary cineangiograms. *Comput. Cardiol.*, pp. 307–310, 1989.

[40] B. Tom, S. Efstratiadis, and A. Katsaggelos. Motion estimation of skeletonized angiographic images using elastic registration. *IEEE Trans. Med.Imaging*, vol. 13, no. 3, pp. 450–460, 1994.

[41] R. Curwen, A. Amini, J. Duncan, and F. Lee. Tracking vascular motion in X-ray image sequences with Kalman snakes. *Comput. Cardiol.*, pp. 109–112, 1994.

[42] M.-P. Dubuisson-Jolly, C. Liang, and A. Gupta. Optimal polyline tracking for artery motion compensation in coronary angiography. *Proc. ICCV*, pp. 414–419, 1998.

[43] T. Saito, M. Misaki, K. Shirato, and T. Takishima. Three-dimensional quantitative coronary angiography. *IEEE Transactions on Biomedical Engineering*, v 37, n 8, p 768-77, Aug. 1990.

[44] M. Garreau, J. Puentes, H. Lebreton, and C. Roux. Understanding coronary artery movement: a knowledge-based approach. *Artificial Intelligence in Medicine*, v 13, n 3, p 207-37, July 1998.

[45] J.L. Coatrieux , J. Puentes, M. Garreau, C. Roux. Towards dynamic cardiac scenes interpretation based on spatial-temporal knowledge. *Artificial Intelligence in Medicine.* Volume 19, Issue 2, Pages 155-183. June 2000.

[46] S. Lu and S. Eiho. Automatic detection of the coronary arterial contours with sub-branches from an x-ray angiogram. *IEEE Computers in Cardiology*, pp. 575-578, 1993.

[47] Y. Tolias and S.M. Panas. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE Trans. on Med. Img.*, vol. 17, pp. 263-273, April 1998.

[48] L. Lecornu, C. Roux, and J.J. Jacq. Extraction of vessel contours in angiograms by simultaneous tracking of the two edges. *IEEE Conf. Eng. in Medicine and Bio.*, vol. 1, pp. 678-679, 1994.

[49] C. Bellemare and J. Meunier. Structural method for tracking coronary arteries in coronary cineangiograms. *Proceedings of the First Canadian Conference on Computer and Robot Vision.* 2004.

[50] H. Zhu and M.H. Friedman. Vessel Tracking by Template String in Angiography. *Medical Image Acquisition and Processing, Proc. SPIE Conf. Multispectral Image Processing and Pattern Recognition.* Wuhan, China, pp. 29-33. October 2001.

[51] S.A.M. Baert, M.A. Viergever, W.J. Niessen. Guide wire tracking in endovascular interventions. *IEEE Transactions on Medical Imaging*, vol. 22, pp. 965-972. 2003.

[52] J. Meunier, M. Bourassa, M. Bertrand, M. Verreault, and G. Mailloux. Regional epicardial dynamics computed from coronary cineangiograms. *Comput. Cardiol.* pp. 307–310. 1989.

[53] G. Hamarneh, K. Althoff, T. Gustavsson. Snake Deformations Based on Optical Flow Forces for Contrast Agent Tracking in Echocardiography. *Swedish Symposium on Image Analysis*, pp. 45-48.2000.

[54] S. Ruan, A. Bruno, R. Collorec, and J. Coatrieux. 3-D motion and reconstruction of coronary networks. *Proc. IEEE EMBS*, vol. 5, pp. 2048–2049. 1992.

[55] C. Cañero, P. Radeva, R. Toledo, J. Villanueva, and J. Mauri. 3-D curve reconstruction by biplane snakes. *Proc. ICPR*, vol. 4, pp. 563–566. 2000.

[56] G. Shechter, F. Devernay, E. Coste-Manière, A. Quyyumi, and E. R. McVeigh. Three-Dimensional Motion Tracking of Coronary Arteries in Biplane Cineangiograms. *IEEE Transactions on Medical Imaging,* Vol. 22, No. 4, pp. 493-503. April 2003.

[57] Wollschlaeger H, Lee P, Zeiher A, Solzbach U, Bonzel T, Just H. Mathematical tools for spatial computations with biplane isocentric X-ray equipment. *Biomedizinische Technik*; 31: 101-106,1986.

[58] Metz CE, Fencil LE. Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views. *Med Phys*; 16: 45-51,1989.

[59] Hoffmann KR, Metz CE, Chen Y. Determination of 3D imaging geometry and object configurations from two biplane views: an enhancement of the Metz-Fencil technique. *Med Phys*; 22: 1219-1227,1995.

[60] W. Faig. Calibration of close-range photogrammetry systems: Mathematical formulation. *Photogrammetric Engineering and Remote Sensing*, 41(12):1479–1486, 1975.

[61] I. Shimizu, Z. Zhang, S. Akamatsu, and K. Deguchi. Head pose determination from one image using a generic model. *Proceedings of the IEEE Third International Conference on Automatic Face and Gesture Recognition*, pages 100–105, Nara, Japan, Apr. 1998.

[62] O. Faugeras and G. Toscani. Camera Calibration for 3D Computer Vision. *International Workshop on Machine Vision and Machine Intelligence*, pp. 240-247, Tokyo, 1987.

[63] J. Weng, P. Cohen and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(10):965-980, 1992.

[64] R. Willson, *Modeling and Calibration of Automated Zoom Lenses*, Ph.D. thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, January 1994.

[65] R. Lenz and R. Tsai. Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:713-720, 1988.

[66] B. Triggs. Autocalibration from planar scenes. *Proceedings of the 5th European Conference on Computer Vision*, pages 89–105, Freiburg, Germany, June 1998.

[67] S. J. Maybank and O. D. Faugeras. A theory of self-calibration of a moving camera. *The International Journal of Computer Vision*, 8(2):123–152, Aug. 1992.

[68] R. Tsai. An efficient and accurate camera calibration technique for 3D machine vision. *Proc. Computer Vision and Pattern Recognition*, 1986.

[69] R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323-331, August 1987.

[70] G. Golub and C. van Loan. Matrix Computations. The John Hopkins University Press, Baltimore, Maryland, 3 edition, 1996.

[71] T, Moons, L. Van Gool, M. Proesmans and E. Pauwels, ``Affine reconstruction from perspective image pairs with a relative object-camera translation in between", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no.1, pp. 77-83, Jan. 1996.

[72] R. Hartley. Self-calibration from multiple views with a rotating camera. *Lecture Notes in Computer Science*, Vol. 800-801, Springer-Verlag, pp. 471-478, 1994.

[73] M. Armstrong, A. Zisserman and R. Hartley. Euclidean Reconstruction from Image Triplets. *Computer Vision - ECCV'96*, Lecture Notes in Computer Science, Vol. 1064, Springer-Verlag, pp. 3-16, 1996.

[74] M. Pollefeys and L. Van Gool. A stratified approach to self-calibration. *Proc. 1997 Conference on Computer Vision and Pattern Recognition*, IEEE Computer Soc. Press, pp. 407-412, 1997.

[75] D. Bondyfalat and S. Bougnoux. Imposing Euclidean Constraints During Self-Calibration Processes. *Proc. SMILE Workshop (post-ECCV'98)*, Lecture Notes in Computer Science, Vol. 1506, Springer-Verlag, pp.224-235, 1998.

[76] D. Liebowitz and A. Zisserman. Combining Scene and Auto-calibration Constraints. *Proc. ICCV*, pp.293-300, 1999.

[77] P. Sturm. Critical Motion Sequences for Monocular Self-Calibration and Uncalibrated Euclidean Reconstruction. *Proc. 1997 Conference on Computer Vision and Pattern Recognition*, IEEE Computer Soc. Press, pp. 1100-1105, 1997.

[78] A. Wahle, E. Wellnhofer, I. Mugaragu, H.U. Sauer, H. Oswald and E. Fleck. Assessment of diffuse coronary artery disease by quantitative analysis of coronary morphology based upon 3-D reconstruction from biplane angiograms. *IEEE Transactions on Medical Imaging* 14: 230-241.1995.

[79] A. Wahle. PraÈzise dreidimensionale Rekonstruktion von GefaÈ ûsystemen aus biplanen angiographischen Projektionen und deren klinische Anwendung. No. 152 in *Fortschritt Berichte, Reihe Biotechnik* (17), DuÈ sseldorf: VDI Verlag1997.

[80] M.J. Potel, J.M. Rubin, S.A. MacKay, A.M. Aisen, J. Al-Sadir and R.E. Sayre. Methods for evaluating cardiac wall motion in three dimensions using bifurcation points of the coronary arterial tree. *Invest Radiol* 8: 47-57.1983.

[81] K.R. Hoffmann, J. Esthappan, S. Li and C.A. Pelizzari. A simple technique for calibrating imaging geometries. *Proc SPIE* 2708: 371-375.1996.

[82] S.Y.J. Chen and C.E. Metz. Improved determination of biplane imaging geometry from two projection images and its application to three-dimensional reconstruction of coronary arterial trees. *Med Phys* 24: 633- 654.1997.

[83] C.J. Henri and T.M. Peters. Three dimensional reconstruction of vascular trees. Theory and methodology. *Med Phys* 23: 197-204. 1996.

[84] R. Close, C. Morioka and J.S.Whiting. Automatic correction of biplane projection imaging geometry. *Med Phys* 23: 133-139.1996.

[85] J. Woodward. Change Detection in MRI Brain Scan Data. MSc thesis, Dept. Artificial Intelligence, Univ. of Edinburgh, 1990.

[86] A. Goshtasby, D. Turner, and L. Ackerman. Matching of Tomographic Slices for Interpolation. *IEEE Trans. Medical Imaging*, vol. 11, no. 4, pp. 507-516, 1992.

[87] A.E. Johnson and S. Kang. Registration and Integration of Textured 3D Data. *Proc. Int'l Conf. Recent Advances in 3D Digital Imaging and Modeling*, pp. 234-241, 1997.

[88] R. Bolles and R. Cain. Recognising and Locating Partially Visible Objects: The Local-Feature-Focus Method. *Int'l J. Robotics Research* vol. 1, pp. 57-82, 1982.

[89] F. Verbeek. Three-Dimensional Reconstruction of Biological Objects from Serial Sections Including Deformation Correction. PhD thesis, Technische Universiteit Delft, 1995.

[90] G. Subsol, J.-P. Thirion, and N. Ayache. A Scheme for Automatically Building Three-Dimensional Morphometric Anatomical Atlases: Application to a Skull Atlas. *Medical Image Analysis*, vol. 2, no. 1, pp. 37-60, 1998.

[91] H. Chui, J. Rambo, J. Duncan, R. Schultz, and A. Rangarajan. Registration of Cortical Anatomical Structures via Robust 3D Point Matching. *Proc. Int'l Conf. Information Processing in Medical Imaging*, pp. 168-181, 1999.

[92] R. Feris, R. Raskar, C. Longbin, T. Kar-Han and M. Turk. Multiflash Stereopsis: Depth-Edge-Preserving Stereo with Small Baseline Illumination. *Pattern Analysis and Machine Intelligence, IEEE Transactions on.* Volume 30, Issue 1, Page(s):147-159. Jan. 2008.

[93] R. Hartley. In Defense of the 8-Point Algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 580-593.1997.

[94] Z. Zhang Z. Determining the Epipolar Geometry and Its Uncertainty: A Review Two-Body Segmentation from Two Perspective Views. *Int'l J. Computer Vision*, vol. 27, no. 2, pp. 161-195. 1998.

[95] P.H.S. Torr and D.W. Murray. The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix. *Int'l J. Computer Vision*, vol. 24, no. 3, pp. 271-300. 1997.

[96] G. Shechter. Respiratory Motion of the Heart: Implications for Magnetic Resonance Coronary Angiography. PhD Dissertation. 2003.

[97] K. Sprague, M. Drangova and G. Lehmann. Coronary x-ray angiographic reconstruction and image orientation. *Medical Physics*; Journal Volume: 33; Journal Issue: 3; p.707-718. 2006.

[98] N. Sang, W. Peng, H. Li H, Z. Zhang and T. Zhang. 3D reconstruction of the coronary tree from two Xray angiographic views. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, v 6144 III, Medical Imaging 2006: Image Processing, pp 61444Y. 2006.

[99] C. Blondel, B. Malandain, R. Vaillant and N. Ayache. Reconstruction of Coronary Arteries from One Rotational X-Ray Projection Sequence. *Medical Imaging, IEEE Transactions* on Volume 25, Issue 5, pp 653-663. 2004.

[100] C. Venaille, D. Mischler, R. Collorec, J. Y. Catros, and J. L. Coatrieux. Automatic 3D reconstruction of vascular networks from three projections: a simulated annealing approach. *Proceedings, 11th Annual Intl. Conf. of the IEEE Engr. in Medicine and Biology Society*, pages 565–566, Seattle, November 1989.

[101] F. Cheriet, J. Meunier, J. Lesperance and M. Bertrand. 3D motion and structure estimation using temporal and stereoscopic point matching in biplane cineangiography. *Computers in Cardiology*. 8-11 Sept. Page(s): 409-412. 1996.

[102] M. Pollefeys and J. Repko. 3D Models from Extended Uncalibrated Video Sequences: Addressing Key-frame Selection and Projective Drift. *Proc. 3DIM*. 2005.

[103] D. Tzovaras.; N. Ploskas and M.G. Strintizis. Rigid 3-D motion estimation using neural networks and initially estimated 2-D motion data. *Circuits and Systems for Video Technology, IEEE Transactions on*. Volume 10, Issue 1, Page(s):158 – 165. Feb. 2000.

[104] E. Delage, H. Lee and Y.Ng. A dynamic Bayesian network model for autonomous 3d reconstruction from a single indoor image. *CVPR*. 2006.

[105] R. Cipolla, G. Fletcher, and P. Giblin. Surface geometry from cusps of apparent contours. *ICCV*, 1995.

[106] A. Criminisi, I. Reid, and A. Zisserman. Single view metrology. *IJCV*, 40(2), Nov. 2000.

[107] F. Han and S.-C. Zhu. Bayesian reconstruction of 3D shapes and scenes from a single image. *Workshop on Higher- Level Knowledge in 3D Modeling and Motion Analysis*, 2003.

[108] B. Horn. Obtaining Shape from Shading Information. McGraw-Hill, 1975.

[109] A. Witkin. Recovering surface shape and orientation from texture. *AI*, 17(1–3):17–45, 1981.

[110] D. Hoiem, A. Efros, and M. Hebert. Automatic photo popup. *SIGGRAPH*, 2005.

[111] D. Hoiem, A. Efros, and M. Hebert. Geometric context from a single image. *ICCV*, 2005.

[112] J. Atick, P. Griffin, and A. Redlich. Statistical approach to shape from shading: Reconstruction of three-dimensional face surfaces from single two-dimensional images. *Neural Computation*, 8(6):1321–1340, 1996.

[113] V. Blanz and T. Vetter. A morphable model for the synthesis of 3D faces. *SIGGRAPH*, 1999.

[114] R. Dovgard and R. Basri. Statistical symmetric shape from shading for 3D structure recovery of faces. *ECCV*, 2004.

[115] S. Romdhani and T. Vetter. Efficient, robust and accurate fitting of a 3D morphable model. *ICCV*, 2003.

[116] T.Hassner and R. Basri. Example Based 3D Reconstruction from Single 2D Images Beyond Patches. *Workshop at IEEE Conference on Computer Vision and Pattern Recognition.* June 2006.

[117] D. Sherknies, J. Meunier, R. Mongrain, et J. C. Tardif. 3D trajectory assessment of an IVUS transducer from single-plane cineangiograms: a phantom study. *IEEE Transactions On Biomedical Engineering*. September 2004.

[118] M. Jourdain et al. 3D reconstruction of an IVUS transducer trajectory with a single view in cineangiography. *SPIE Medical Imaging*. pp.1430-1437. April 2005.

[119] P. Fallavollita, P. Savard and G. Sierra. Fluoroscopic navigation to guide RF catheter ablation of cardiac arrhythmias. EMBS 26th Annual International Conference. Volume 1, Page(s):1929 – 1932. 2004.

[120] P. Fallavollita. Fluoroscopic navigation to guide RF catheter ablation of cardiac arrhythmias. École Polytechnique, Montreal (Canada), 75 pages, 2004.

[121] Homomorphic filtering notes. Available at: http://www.cs.sfu.ca/~stella/papers/blairthesis/main/node35.html.Accessed July 1, 2008

[122] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7):629-639, 1990.

[123] G. Gilboa, N. Sochen, and Y. Zeevi. Regularized shock filters and complex diffusion. In Proceedings of ECCV, pages 399-413, Copenhagen/Denmark, 2002.

[124] Morphological filtering notes. Available at: http://homepages.inf.ed.ac.uk/rbf/HIPR2/strctel.htm. Accessed July 21, 2008.

[125] J. Y. Bouguet. Pyramidal Implementation of the Lucas Kanade Feature Tracker. Intel Open Source Computer Vision Library (OpenCV) distribution. 2004.

[126] C. Xu, D.L. Pham, and J.L. Prince. Medical Image Segmentation Using Deformable Models. Chapter 3, pp. 129-174, *SPIE Press*, 2000.

[127] G.L. Mariottini and D. Prattichizzo. EGT for Multiple View Geometry and Visual Servoing. Robotics and Vision with Pinhole and Panoramic Cameras. IEEE Robotics and Automation Magazine, 12(4):26--39, December 2005.

[128] G. Li and S. Zucker. Contextual Inference in Contour-Based Stereo Correspondence. *International Journal of Computer Vision (IJCV),* 69(1):59-75. 2006.

[129] R. Cipolla and A. Zisserman. Qualitative surface shape from deformation of image curves. *International Journal of Computer Vision*, 8:53–69,1992.

[130] R. Cipolla and P. Giblin. Visual Motion of Curves and Surfaces. Cambridge Univ. Press. 2000.

[131] D. Zikic, M. Groher, A. Khamene, N. Navab. Deformable Registration of 3D Vessel Structures to a Single Projection Image. *SPIE Medical Imaging*, San Diego, California, USA, 16-21 February 2008.

[132] DW. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of Society for Industrial and Applied Mathem*atics. 11(2):431–441. 1963.

[133] S. A. Stevens. Clinical data from The International Heart Institute of Montana. Available at: http://vortex.bd.psu.edu/~sas56/Research/. Accessed July 1, 2008.

[134] Y. Chunke, S. Oe and K.Terada. Estimation of three-dimensional motion information from optical flow using subspace method. *Transactions of the Institute of Electrical Engineers of Japan.* Part C, v 121-C, n 7, p. 1187-94. July 2001.

[135] C. Li, C.Y. Kao, J.C. Gore and Z Ding. Implicit Active Contours Driven by Local Binary Fitting Energy. *CVPR*. 2007

Ph.D PUBLICATIONS

Journal Articles

P. Fallavollita, F. Cheriet, P. Savard. Single view depth estimation of ablation catheters for arrhythmia procedures. April 2008. (ready for submission)

P. Fallavollita and F. Cheriet. Optimal 3D Reconstruction of Coronary Arteries for 3D Clinical Assessment. October 2007. *Computerized Medical Imaging and Graphics.* (Accepted, in print)

Conference Proceedings

P. Fallavollita, F. Cheriet. Robust Coronary Artery Tracking from Fluoroscopic Image Sequences. 2007. *Lecture Notes in Computer Science 4633- International Conference on Image Analysis and Recognition*, pp. 889-898.

P. Fallavollita and F. Cheriet. Towards an Automatic Coronary Artery Segmentation Algorithm. 2006. *Engineering in Medicine and Biology Society, Proceedings of the 28th IEEE EMBS Annual International Conference*, pp. 3037-3040.

P. Fallavollita and F. Cheriet. Vers un algorithme automatique de segmentations des artères coronaires. 2006. *74th Congress of ACFAS*, McGill University.